

ESSAYS IN PRODUCTION SCHEDULING
WITH JUST-IN-TIME RELATED PERFORMANCE MEASURES

BY
FRANZ-JOSEF KRÄMER

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1994

ACKNOWLEDGMENTS

The author gratefully acknowledges Dr. Chung-Yee Lee, the committee chairman, and Dr. Sherman Bai, the cochairman, for their expert guidance and their constant encouragement at any time throughout this study. Appreciation is expressed to the Department of Industrial and Systems Engineering of the University of Florida and in particular to its chairman Dr. Jack Elzinga for the financial support.

The author wishes to thank his family and friends for their continuous support during the course of this work. Special thanks are due to his deceased father who in his early years created and nurtured his commitment to education and who has made abundant provision for his family, and to his loving mother who let him go across the Atlantic to pursue his goals and dreams.

The author also would like to thank God. So many times during the creative part of this research, the Lord has proven Himself faithful to His promise:

Commit your works to the Lord,
and your thoughts shall be established.

(Proverbs 16:3).

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
ABSTRACT	v
1 INTRODUCTION	1
2 COMMON DUE-WINDOW SCHEDULING	8
2.1 Introduction	8
2.2 Notation and General Properties	9
2.3 Decision Variable Case	12
2.4 Parameter Case	18
2.5 Final Remarks and Conclusions	23
3 PARALLEL DUE WINDOW SCHEDULING	24
3.1 Introduction	24
3.2 Notation and General Properties	26
3.3 Complexity and Optimality Properties	29
3.4 The Two Machine Case	38
3.5 Generalization for the m machine case	49
3.6 Summary	53
4 OPTIMAL CONTROL OF A PRODUCTION SYSTEM WITH PERIODIC MAINTENANCE	54
4.1 Introduction	54
4.2 Problem Definition	57
4.3 A One-Cycle Problem and its Optimal Solution	59
4.3.1 Singular Intervals	62
4.3.2 Switching Times and Switching Points	68
4.3.3 Initial Control Level $u^*(0)$	73
4.3.4 Optimal Solution	78
4.4 A Two-Cycle Problem and its Optimal Solution	111
4.4.1 Singular Intervals for Problem P_2	114
4.4.2 Switching Points for Problem P_2	115
4.4.3 Initial Control Level for Problem P_2	116
4.4.4 Optimal Solution for Problem P_2	116
4.5 Optimal solution to the N -cycle problem ($N \geq 3$)	142
4.5.1 Singular Intervals for Problem P_N	144
4.5.2 Switching Points for Problem P_N	144
4.5.3 Initial Control Level for Problem P_N	144
4.5.4 Optimal Solution for Problem P_N	144

4.6	Transient State and Steady State	155
4.7	Summary	158
APPENDICES		
A	PROOF OF LEMMA 1	160
B	ALGORITHM FOR $m = 2$ PARALLEL MACHINE CASE	165
C	CALCULATION FOR A PRODUCTION SYSTEM WITH PERIODIC MAINTENANCE	170
C.1	The range of $S_1(x(0))$ for Problem P_1 (Case B4-2)	170
C.2	Switching Point for Problem P_1 (Case B4-2)	172
C.3	The Switching Curve	173
C.4	The First Derivative of $J_1^*(x(0))$	177
C.5	The Derivative of $J_1^*(x(0))$ (Case E4)	178
C.6	The Second Derivative of $J_1^*(x(0))$	181
C.7	A Property of $\frac{d}{dx(0)}J_1^*(x(0))$	183
C.8	A Property of $\frac{d}{dx(0)}J_2^*(x(0))$	187
C.9	An Initial Production Control for Problem P_2	189
C.10	Rewriting a quadratic equation (Case C4-2)	191
C.11	$S_2(x(0))$ is monotone decreasing (C4-2-2)	195
C.12	Evaluation of $S_2(d(a - \frac{1}{2} \frac{db}{U-d}))$ (Case C4-2-2)	197
REFERENCES		199
BIOGRAPHICAL SKETCH		207

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

ESSAYS IN PRODUCTION SCHEDULING
WITH JUST-IN-TIME RELATED PERFORMANCE MEASURES

By

Franz-Josef Krämer

August 1994

Chairman: Dr. Chung-Yee Lee
Cochairman: Dr. Sherman Bai
Major Department: Industrial and Systems Engineering

Just-In-Time concepts are of increasing importance in today's competitive manufacturing environment. In this context, we study several *discrete scheduling problems* as well as a *production flow control problem*.

The *discrete scheduling problems* consider the case where we have a common due window. Jobs which are completed within the due window incur no penalty. Completion before the due window incurs earliness penalty and completion after the due window incurs tardiness penalty. If the location of the due window is a decision variable, we can provide a polynomial algorithm for the single machine case. For two parallel machines, we show that the problem is NP-complete and use dynamic programming to solve it. We present a heuristic with polynomial complexity for m parallel machines and show that the error bound vanishes as the number of jobs increases. If the location of the due window is a given parameter, the

problem is NP-complete even for the single machine case. We use dynamic programming to solve this problem. The main contribution of this research is the extension of the due-date concept to the due-window concept. No prior research has analyzed earliness and tardiness performance measures in combination with due windows.

For the *production flow control problem*, we consider a system consisting of one machine for which maintenance is performed on a periodic basis. The time is continuous and the time horizon is finite. For each cycle, both the time interval of operation and the maintenance interval are deterministic. During the maintenance interval, the system is shut down and cannot produce. One part-type is produced and the demand rate is assumed to be constant. The supply is assumed to be unlimited. In order to make on-time delivery, the objective is to keep the production as close to the demand as possible. However, the maintenance disruptions make the production deviate from the demand. We model the problem as an optimal control problem and solve it by Pontryagin's Minimum Principle. We discuss both the transient state and the steady state. The main contribution is the analysis of the transient state.

CHAPTER 1

INTRODUCTION

In the 1990s, international competition will increase due to a number of reasons. One of these reasons is that advancements in information and process technologies occur at an accelerated pace. Also, the consumer demands are increasingly sophisticated and individualized. This results in the need for operational models which reflect long-term and nonfinancial performance measures such as customer satisfaction with both the product and its service. Revolutionary changes in the management of production systems are yet another factor which contributes to an increased competition.

World-Class-Manufacturing requires an integrated manufacturing strategy, encompassing Total Quality Control (TQC), Just-In-Time (JIT) Production, Factory Automation (FA), and Total Productive Maintenance (TPM). In this dissertation, we will focus our attention to the Just-In-Time production strategy. According to the Just-In-Time philosophy, the ideal situation is to eliminate all inventory. Inventory in its enlarged definition not only refers to either partially or fully completed products, but includes for example also excess amount of labor and machine time which is commonly referred to as idle time. The Just-In-Time philosophy seeks to avoid inventory of any kind. Another motivation to reduce inventory is that inventory hides root causes of a problem. For example, inventory hides problems such as rework, machine downtime, change order, long setups, and vendor delinquencies. It is therefore desirable to identify and eliminate waste components such as overproduction, waiting time, transportation, processing, inventory, movement, and defective products. These are important steps toward a Just-In-Time environment. The process

of creating a Just-In-Time production environment is multi-faceted and involves many concepts. A very important component in this process is the utilization of Just-In-Time oriented production scheduling techniques.

At this point, we note that the area of both sequencing and scheduling is of great importance and is motivated by problems that arise in production planning and control, and in general all situations in which scarce resources have to be allocated to activities over time. General references on the subject are the classic book by Conway, Maxwell and Miller [30], the introductory texts by Baker [17] and French [40], the volume edited by Dempster, Lenstra, and Rinnooy Kan [34], and survey papers such as by Graves [51] in the area of production planning, and by Lawler, Lenstra, Rinnooy Kan, and Shmoys [75] in the general area of Sequencing and Scheduling. *Sequencing and Scheduling* “is concerned with the *optimal allocation of scarce resources to activities over time*” ([75]), and more than any other area in operations research, the theory of sequencing and scheduling is characterized by a virtually unlimited number of problem types. In regard to the deterministic models, the problems can be classified according to their complexity. Complexity theory provides a mathematical framework which can classify problems as “easy” or as “hard” (or NP-complete). Pioneering work in the theory of NP-completeness has been done by Cook [28] and Karp [64] and [65]. The text by Garey and Johnson [46] is a classical guide in this area.

For many years, scheduling research focused on single performance measures that are nondecreasing in job completion times. These are also referred to as regular performance measures. Examples include mean flowtime, mean lateness, number of tardy jobs, and total tardiness. Regular due-date-related performance measures are unable to capture any penalty for early job completion. We thus are required to use nonregular performance measures in order to model the objective of Just-In-Time production scheduling problems. Such performance measures penalize both earliness and tardiness where earliness occurs when a job is completed before the due date and tardiness is incurred when a job is completed after the due date.

In this study, we examine several production scheduling problems where both early and tardy job completions are penalized. The problems can be categorized in two groups: The first group consists of problems which are modelled by *discrete scheduling models*. The second group consists of problems which are modelled by *production flow control models*.

For *discrete scheduling models*, we consider two problems in Chapters 2 and 3. The performance measure is due-date oriented and penalizes both earliness and tardiness. According to the remark above, such a performance measure is a nonregular performance measure. More specifically, we consider a due-date window which consists of an earliest due-date and a latest due-date. There are very few references which allow for a zero penalty when the completion time falls within a time interval (see Chapter 2), and none of them develop the concept of due-window scheduling. A survey of earliness and tardiness related literature is given in Baker and Scudder [20] where the due-date related performance measures in most cases refer to the same point in time for the earliness penalty as well as for the tardiness penalty. While this model is applicable in many situations (for example, a production line where deviation from a point in time affects the processing of the part at the next machine), there are other situations where there is a time window during which no cost is incurred. Consider, for example, the situation where a truck can be loaded overnight. So there is a time window from the time when the driver delivers the truck in the evening to the time when the driver starts the trip the next morning where there is no penalty for the products which are completed and loaded during the night shift. Products which are completed early require holding and products which are finished tardy require special shipment, for example. Thus, for situations where the due-date as one point in time does not satisfactorily reflect the situation, we consider an interval of time for each job during which no penalty is incurred. We refer to this time window as due window.

In the second chapter, we consider a single machine problem where n jobs have to be processed. We have a common due window $[e, d]$ with earliest due-date e and latest due-date d which is the same for each job. For a completion before the earliest due-date e , an

earliness cost is incurred. Completion after the latest due-date d results in tardiness cost. The size of the due window $d - e$ is a given parameter. Jobs which are completed within the due window do not incur a penalty. Our objective is to find a schedule which minimizes the total penalty. We distinguish between the *parameter case*, in which the location of the due window is given, and the *decision variable case*. In the decision variable case, the location of the due window is to be determined. We do not consider cost for due-date assignment.

For the decision variable case and a certain cost structure, we are able to provide an algorithm which requires $O(n)$ time, provided the indices are sorted according to the shortest processing time order. If the indices are not sorted, the complexity is $O(n \log(n))$. Since even the easiest case with unit weights is NP-complete for $d = e$ (see Chapter 2), there is no realistic hope in finding a polynomial algorithm for this problem. Thus, we use dynamic programming to solve the parameter case. We observe that in the case where the size of the due window exceeds the largest processing time, we obtain a more efficient algorithm than for the corresponding problem for the common due date case $d = e$. The results of this research are to appear in Krämer and Lee [72].

In the third chapter, we study the extension of this problem to an environment with parallel machines. Concerning the location of the due window, we consider the decision variable case. We show that the easiest two-machine problem (which has unit cost) is NP-complete and provide a dynamic programming algorithm to solve for the optimal solution. Subsequently, we present a heuristic for the two-machine and for the m -machine case. We give an expression for the upper bound of the absolute error and show that the relative error vanishes as the number of jobs increases. The results of this research are to appear in Krämer and Lee [73]. Chapter 3 concludes the problems with *discrete scheduling models*.

We then turn to the second group of problems which are modelled as *production flow control models*. The term *production flow control model* includes two concepts: *production flow* and *control*. By modelling the production as a *production flow*, we model a discrete process by a continuous production rate. This has the disadvantage of introducing some

inaccuracy into the model and it has the advantage to simplify the analysis. The concept of *control* concerns the control of these production rates. When such decisions are sufficiently less frequent than the event of operation completion, it is reasonable to model a problem in such a way. The problems which are modelled by a production flow control model consider part-types for which the demand rate is assumed to be constant.

Since the early 1980s, there has been an increased interest in developing production flow control models for systems with unreliable machines. Kimemia and Gershwin [66] have studied a flexible manufacturing system of unreliable machines where they used a production flow model. The complicating factor is the unreliability of the machines involved. These break down in a random fashion and the required repair time is also modelled by a random variable. Their main contribution is the qualitative characterization of an optimal control structure. They introduced the notion of a *hedging point*. The hedging point is a production surplus level which is utilized to hedge against random machine failures. The control policy, also referred to as hedging policy or threshold policy, operates as follows: There is a target inventory level (referred to as threshold level or hedging level), and the system tries to reach this level as soon as possible. When the actual inventory is below the target level, production is at full capacity. When the actual inventory is above the target, the machine is shut down. When the target inventory level is reached, the production rate equals the demand. The target inventory level is used to hedge against machine breakdowns. Further, based on assumptions concerning the frequency of events (demand changes occur less frequent than machine failures and operation times are small compared to the time between failures), they have proposed a multi-level hierarchical control algorithm. The top level is used for long-term decisions such as system configuration, the next level is used to control the production flow and the routing. At the lowest level of control are part-dispatch algorithms.

Subsequently, the concept of modeling production as continuous flow, the hedging-point concept, and the concept of a hierarchical control algorithm have been refined and applied to similar problem settings:

Gershwin, Akella, and Choong [49] have improved the algorithms for the various levels of the hierarchical control. Akella, Choong, and Gershwin [4] have studied the performance of this model for an IBM printed circuit card assembly line. They have concluded that a hierarchically structured policy such as given in [49] is very effective in scheduling a FMS and that it can achieve high output with low WIP and can cope with changes and disturbances. Recently, Bai and Gershwin have studied finite buffer capacity models. In [14], they have considered one-part-type systems, multiple-part-type systems in [15], and reentrant systems in [16]. Bai and Burhanpurwala [12] and Bai and Elhafi [13] have studied unreliable systems with non-negligible setup times.

It is difficult to obtain an optimal control for a system which consists of unreliable machines. In the following case, researchers have succeeded: Akella and Kumar [5] considered a manufacturing system consisting of one unreliable machine which produces one part-type. They analytically derive an exact solution and show that the optimal control policy is a hedging policy.

In Chapter 4, we consider a deterministic *production flow control model* which can be viewed as an extreme case of a stochastic model where the variance of the time between failures and the time between repairs is zero, i.e., we have a constant interval of time for the operation and a constant interval of time set aside for the maintenance (repair) of the machine. Thus, the production system under consideration consists of one machine and maintenance is performed on a periodical (or cyclical) basis. Within a cycle, the operational time is a constant. The time for maintenance is a constant as well. The production system produces one part-type, for which the demand rate is assumed to be constant. We formulate the problem as an optimal control problem. The methodology used to solve the problem is Pontryagin's Minimum Principle as found in Pontryagin et al. [86] or in Kirk [67].

We first examine the structure of the optimal control during the time interval of operation and we determine the initial control level according to the initial production surplus

level. We then proceed to formulate and solve the problem for one cycle. Valuable insights are gained concerning the solution structure. The results of the one-cycle problem are available in Krämer and Bai [70].

Next, we proceed to analyze the two-cycle problem. Subsequently, we solve the general N -cycle problem. It is noteworthy to observe that the optimal control during the first cycle of the N -cycle problem for $N \geq 3$ is identical to the optimal control during the first cycle of the two-cycle problem. For our problem, we define what is understood by the term steady state and express the number of cycles to reach this state as a function of the initial production surplus. The results for the problem with $N \geq 2$ cycles and the discussion of the steady state are available in Krämer and Bai [71].

CHAPTER 2

COMMON DUE-WINDOW SCHEDULING

2.1 Introduction

Within the area of production scheduling, it becomes increasingly important not only to produce high quality products, but to deliver them at the right time. The right time not only requires to meet the customer's due date, but due to inventory costs, it also requires to avoid earliness. This is reflected by the growing importance of Just-In-Time production. The performance measure which is associated with the Just-In-Time concept is nonregular (see Baker [17]). A vast amount of literature has accumulated in the area of earliness and tardiness penalties, especially in the context of common due-date scheduling. For an excellent review of earliness-tardiness problems, see Baker and Scudder [20]. Hall and Posner [55] and Hall, Kubiak, and Sethi [56] provided the first two pseudopolynomial algorithms to solve the problem. Related literature includes Lee, Danusaputro, and Lin [77], and Hoogeveen and Van De Velde [58]. Liman and Lee [80] provided the first finite error bound for the heuristic method. Federgruen and Mosheiov [39] extended the objective to certain nonlinear objective functions.

The Just-In-Time Scheduling model assumes that there is only one acceptable completion time, and any earliness or tardiness is penalized. This model fits intra-manufacturing systems situations like, for example, a production line with adjacent machines processing a product in sequence. If no buffer is used, starvation or blockage will occur when the production is not synchronized. A buffer could be used at additional cost. This type of situation perfectly fits the traditional due-date models (due date as point-in-time).

However, manufacturers often consider a due date as an interval rather than a point in time. Thus, for each job there are earliest and latest due dates. Any job finished after its

latest due date is considered tardy, and no job can be delivered before its earliest due date. If it finished earlier, a job must be held until its earliest due date thus incurring a holding cost. The period between its earliest and latest due date is the due window. A job finished within its due window incurs no penalty (Lee [76]).

Anger, Lee, and Martin-Vega [6] formally introduced the term *due window* and provided a polynomial algorithm to minimize the number of jobs which are completed outside their due windows, where the size of the largest window is less than or equal to the smallest processing time.

Sidney [102] considered earliness and tardiness penalties with target starting times and due dates, and implicitly used the idea of due windows. Lakshminarayan et al. [74] developed a more efficient procedure for the same problem. Cheng [26] studied the common due-window problem where the size of the window was constrained to be less than the shortest processing time, and the penalty was calculated as the absolute deviation from the middle of the due window. Drawing on his practical experience, Lee [76] has modeled and solved several constant-window scheduling problems. He has focused on dual criteria scheduling, differentiating between the cases where idle time is allowed and where it is not.

Our purpose is to extend the earliness and tardiness measures from the due-date case to the due-window case. We consider common due windows.

This chapter is organized in the following way: The next section introduces notation and general properties. In Sections 2.3 and 2.4, we consider cases where the due window is a decision variable and where the due window is a parameter, respectively. In Section 2.5, we summarize the results and comment on future research directions.

2.2 Notation and General Properties

We assume all our data to be integral. We have n jobs, and for each job J_j ($j = 1, \dots, n$), we use the following job attributes:

1. Processing time p_j

2. Starting time S_j

3. Completion time C_j

The earliness penalty is α per unit of time and the tardiness penalty is β per unit of time for each job. We assume jobs are indexed in *shortest-processing-time* (SPT) order, i.e., $p_1 \leq p_2 \leq \dots \leq p_n$. For a given sequence, $[i]$ denotes the i^{th} job. If a particular sequence of k jobs satisfies $p_{[1]} \geq p_{[2]} \geq \dots \geq p_{[k]}$, we say the sequence is in *longest-processing-time* (LPT) order. We also refer to the longest processing time as $p_{\max}$, i.e., $p_{\max} = p_n$.

A window is defined as an interval $[e, d]$, where e is the earliest due date and d is the latest due date. The due window is common to all jobs, and we consider its size $K = d - e$ always as a given parameter of the problem with integral value. Jobs which are completed within the due window do not incur a penalty. The earliness and tardiness of a job is defined as $E_j = (e - C_j)^+$ and $T_j = (C_j - d)^+$, respectively, where $x^+ = \max(x, 0)$. The penalty of job J_j is a function of the completion time C_j and is defined to be

$$\alpha(e - C_j)^+ + \beta(C_j - d)^+ = \alpha E_j + \beta T_j.$$

Our purpose is to find a schedule to minimize $Z = \sum_j (\alpha E_j + \beta T_j)$. We distinguish between the *parameter case*, in which the location of the due window is given, and the *decision variable case*. In the decision variable case, we do not consider cost for due-date assignment.

For simplification, we introduce partial sums. $MS = \sum_{j=1}^n p_j$ represents the total processing time of all jobs, $A_l = \sum_{j=1}^l p_j$ the total processing time of jobs $1, \dots, l$. We assume that $MS > d - e$ since otherwise we could schedule all jobs in the due window and there would be no problem to be solved. We denote the cardinality of a finite set S by $n(S)$.

The following properties are well known for the common due date problem, i.e., $d = e$ (Kanet [63], Hall, Kubiak, and Sethi [56]). The first two properties apply both to the decision variable case as well as to the parameter case.

Property 1: An optimal schedule exists without machine idle time between the first and the last job.

Property 2: In any optimal schedule, early jobs ($E_j > 0$) are sequenced in LPT order, tardy jobs ($T_j > 0$) starting after the common due date ($S_j \geq d$) are sequenced in SPT order.

Property 3: If d is a decision variable, then an optimal schedule exists in which $C_j = d$ for some j .

Property 4: If d is a given parameter, then an optimal schedule exists in which $C_j = d$ for some j or the first job starts at zero, i.e., $S_{[1]} = 0$.

We now consider the extension of those properties from the common due-date problem ($d = e$) to the common due-window problem ($e < d$). We can see that the first property also holds for the common due-window case. We adjust the second property for the common due-window (CDW) case as follows:

Property (CDW) 2: For every optimal schedule, jobs completing before e ($C_j \leq e$) are sequenced in LPT order, and jobs starting after d ($S_j \geq d$) are sequenced in SPT order.

The W -jobs are not included in *Property (CDW) 2*: We will refer to them in part (c) of the following lemma which also extends the third and fourth property for the common due-window case:

Lemma 1 (a) *For the common due-window problem, if d is a decision variable, an optimal schedule exists where one job finishes at either e or d :*

either (i) $C_j = d$ for some j

or (ii) $C_j = d - K = e$ for some j

(b) *If d is a given parameter, an optimal schedule exists where at least one of the following conditions will hold:*

- either (i) $C_j = d$ for some j
 or (ii) $C_j = d - K = e$ for some j
 or (iii) $S_{[1]} = 0$

(c) V-shape Property:

- (i) In both parameter and decision variable cases, if there is an optimal schedule with $C_j = e$ for some j , then there is also an optimal schedule with $C_j = e$ for some j which also satisfies that the E-jobs are sequenced in LPT-order and the remaining jobs are sequenced in SPT order. We describe this schedule as having a V-shape around e .
- (ii) In both parameter and decision variable cases, if there is an optimal schedule with $C_j = d$ for some j , then there is also an optimal schedule with $C_j = d$ for some j which also satisfies that the T-jobs are sequenced in SPT-order and the remaining jobs are sequenced in LPT order. We describe this schedule as having a V-shape around d .
- (iii) For the parameter case, if there is an optimal schedule with $S_{[1]} = 0$ and with at least one job completing within the due window ($e \leq C_j \leq d$), then there is an optimal schedule with $S_{[1]} = 0$ which also satisfies that all jobs with $S_j < e$ are sequenced in LPT order and the remaining jobs are sequenced in SPT order.

Proof. See Appendix A.

QED

2.3 Decision Variable Case

Kanet [63] provided an $O(n \log(n))$ algorithm to solve the unit weight case ($\alpha_j = \beta_j = 1$ for all jobs J_j) for the common due date ($d = e$) problem. There are many optimal solutions for this problem. Bagchi, Sullivan, and Chang [10] provided an algorithm to find all the optimal solutions. Lee, Danusaputro, and Lin [77] provided an $O(n^2)$ algorithm for the constant-weight case. Their problem differed from Bagchi, Sullivan, and Chang's [10] in

that Lee, Danusaputro, and Lin included the number of tardy jobs in the objective function. We extend the constant-weight case to the due-window case and provide an $O(n * \log(n))$ algorithm to solve the problem. The main difference between common due-date and common due-window scheduling is the determination of the set of jobs W which are completed within the due window.

The following definitions are used in the decision-variable case.

Definition 1 For a schedule σ , we define the following three sets:

$$E(\sigma) = \{J_j : C_j(\sigma) \leq \epsilon\} \quad (2.1)$$

$$T(\sigma) = \{J_j : S_j(\sigma) \geq d\} \quad (2.2)$$

$$W(\sigma) = \{J_j : C_j(\sigma) > \epsilon \text{ and } S_j(\sigma) < d\} \quad (2.3)$$

If there is no ambiguity, we omit the parameter σ and write E , T , and W instead. Jobs in E are called E-jobs and are completed at or before the common due window. Jobs in T are called T-jobs and start at or after the window. Set W contains the W-jobs. The cardinality of these sets is denoted by $n(E)$, $n(T)$, and $n(W)$.

According to Equation (2.3), the W -jobs cover the due window. Since we consider only schedules σ which have no idle time during the due window, $\sum_{j \in W(\sigma)} p_j \geq d - \epsilon$ holds. This motivates the following definition:

Definition 2 Define the number nw^* as follows:

$$nw^* = \min \left\{ l : \sum_{j=1}^l p_j \geq d - \epsilon \right\} \quad (2.4)$$

To express the objective function in terms of processing times (Lemma 3, below) and to simplify the discussion, we introduce the concept of cumulative weights. According to Lemma 1(a), we only need to consider schedules with $C_j = \epsilon$ or $C_j = d$ for some job J_j .

Definition 3 We define the cumulative weight $cw(J_j, \sigma)$ for job J_j w.r.t. schedule σ as follows:

1. If job J_j is the k^{th} E-job in schedule σ , i.e., $J_j \in E(\sigma)$ and $[k] = j$, we define

$$cw(J_j, \sigma) = (k - 1)\alpha.$$

2. If job J_j is the k^{th} last T-job in schedule σ , i.e., $J_j \in T(\sigma)$ and $[n - k + 1] = j$, we define

$$cw(J_j, \sigma) = k\beta.$$

3. If job J_j is a W-job in schedule σ , i.e., $J_j \in W(\sigma)$, we define

$$cw(J_j, \sigma) = \begin{cases} (n(E(\sigma)))\alpha & \text{if } C_k = d \text{ for some } J_k \\ (n(T(\sigma)) + 1)\beta & \text{otherwise} \end{cases}$$

We also define

$$cw(W, \sigma) = \begin{cases} (n(E(\sigma)))\alpha & \text{if } C_k = d \text{ for some } J_k \\ (n(T(\sigma)) + 1)\beta & \text{otherwise} \end{cases} \quad (2.5)$$

since the cumulative weight of the W-jobs is the same for every W-job.

Using the definitions for E , T , and W , we illustrate the two solution structures according to Lemma 1(a) and show the cumulative weights beneath the indices in Figure 2.1. These will be used in Lemma 3. In Lemma 4, we will show that the W -jobs consist of the nw^* smallest jobs.

Lemma 2 *An optimal schedule σ must satisfy*

$$\max\{(n(E(\sigma)) - 1)\alpha, n(T(\sigma))\beta\} \leq \min\{n(E(\sigma))\alpha, (n(T(\sigma)) + 1)\beta\} \quad (2.6)$$

Proof. According to Lemma 1 (a), we must either have $C_j = e$ or $C_j = d$ for some j . We prove the Lemma by contradiction:

Assume that $n(T(\sigma))\beta > n(E(\sigma))\alpha$ holds. Obtain a schedule σ' by moving the due window one unit to the right. If $C_j = e$ for some j in schedule σ , then we have (whether $C_j = d$ for some j or whether $C_j \neq d$ for all j)

$$cost(\sigma') - cost(\sigma) = n(E(\sigma))\alpha - n(T(\sigma))\beta < 0.$$

				e				d			
job	e_1	$\dots$	$e_{n(E)}$		$w_1, \dots, w_{n(W)}$		$t_{n(T)}$		$\dots$	t_1	
cum. weight	$0 * \alpha$	$\dots$	$(n(E) - 1)\alpha$				$n(T)\beta$		$\dots$	β	

				e				d			
job	e_1	$\dots$	$e_{n(E)}$		$w_1, \dots, w_{n(W)}$		$t_{n(T)}$		$\dots$	t_1	
cum. weight	$0 * \alpha$	$\dots$	$(n(E) - 1)\alpha$				$n(T)\beta$		$\dots$	β	

Figure 2.1. Two solution structures: $C_j = e$ and $C_j = d$ for some j

If $C_j \neq e$ for all j in schedule σ , then we have

$$\text{cost}(\sigma') - \text{cost}(\sigma) = (n(E(\sigma)) - 1)\alpha - n(T(\sigma))\beta < n(E(\sigma))\alpha - n(T(\sigma))\beta < 0.$$

In both cases, schedule σ is non-optimal, a contradiction.

Now assume that $(n(E(\sigma)) - 1)\alpha > (n(T(\sigma)) + 1)\beta$ in schedule σ . Obtain a schedule σ'' by moving the jobs one unit to the right. The proof of $\text{cost}(\sigma'') - \text{cost}(\sigma) < 0$ is symmetric to the case above.

Thus, an optimal schedule σ satisfies both $n(T(\sigma))\beta \leq n(E(\sigma))\alpha$ and $(n(E(\sigma)) - 1)\alpha \leq (n(T(\sigma)) + 1)\beta$. Since we also have $(n(E(\sigma)) - 1)\alpha < n(E(\sigma))\alpha$ and $n(T(\sigma))\beta < (n(T(\sigma)) + 1)\beta$, Equation (2.6) holds. **QED**

Lemma 3 *Let σ be a schedule with no idle time between jobs ($C_{[n]} - S_{[1]} = MS$) where $C_j = e$ or $C_j = d$ for some j . The weighted earliness and tardiness penalty can be expressed by cumulative weights for the processing times as follows:*

$$\sum_{j=1}^n (\alpha E_j(\sigma) + \beta T_j(\sigma)) = \left\{ \sum_{j=1}^n cw(J_j, \sigma) * p_j \right\} - cw(W, \sigma) * (d - e) \quad (2.7)$$

Proof. According to Lemma 1(a), we consider the two cases which $C_j = e$ and $C_j = d$ for some j (Figure 2.1). In Figure 2.1, the E -jobs are indexed $e_1, \dots, e_{n(E)}$, the T -jobs are indexed $t_1, \dots, t_{n(T)}$, and the W -jobs are indexed $w_1, \dots, w_{n(W)}$.

Now consider job J_{e_j} . Each unit of processing time of job J_{e_j} contributes one unit of earliness to each of the jobs $J_{e_1}, \dots, J_{e_{j-1}}$. Thus, job J_{e_j} contributes $((j-1)\alpha)p_j$ towards the objective function which depends on the number of jobs which precede job J_{e_j} and on the processing time of job J_j . Likewise, each unit of processing time of job J_{t_j} contributes one unit of tardiness to each of the jobs $J_{t_1}, \dots, J_{t_j}$. Thus, the contribution of job J_{t_j} towards the objective function is $(j\beta)p_j$.

The portion of W -jobs exceeding the due window is $(\sum_{j=1}^{n(W)} p_j) - (d - \epsilon)$. In case of $C_j = d$ for some j , this portion would have cumulative weight $n(E)\alpha$. Otherwise ($C_j \neq d$ for all J_j), we have $C_j = \epsilon$ for some job J_j and the cumulative weight is $(n(T) + 1)\beta$. According to the definition of $cw(W, \sigma)$, we can write:

$$\begin{aligned} \sum_{j=1}^n (\alpha E_j(\sigma) + \beta T_j(\sigma)) &= \sum_{j=1}^{n(E)} ((j-1) * \alpha) * p_{e_j} + \sum_{j=1}^{n(T)} (j * \beta) * p_{t_j} \\ &+ cw(W, \sigma) * (\sum_{j=1}^{n(W)} p_{w_j} - (d - \epsilon)). \end{aligned}$$

Using the definition of cumulative weights for the jobs, we can write

$$\sum_{j=1}^n (\alpha E_j(\sigma) + \beta T_j(\sigma)) = \{\sum_{j=1}^n cw(J_j, \sigma) * p_j\} - cw(W, \sigma) * (d - \epsilon)$$

QED

Lemma 4 *An optimal schedule σ exists wherein the set of W -jobs is the set of the nw^* smallest jobs, i.e.,*

$$W(\sigma) = \{J_1, \dots, J_{nw^*}\}$$

Proof. See Appendix A.

QED

Remark 1 *In the following algorithm, we first determine the number of W -jobs and then sequence the E -jobs and the T -jobs. We process the longest job (n) first (with cumulative weight 0), process the second longest job ($n-1$) at position 2 or n , depending on $\min(\alpha, \beta)$, and so forth. In general, the next job is processed before or after the due window depending*

on whether the next early job has a better cumulative weight than the next tardy job. Then the W-jobs are sequenced, and finally, we determine the location of the due window.

Algorithm 1

1. Determine the number nw^* :

$$nw^* = \min\{l : \sum_{j=1}^l p_j \geq d - \epsilon\}$$

2. Sequence jobs $n, \dots, nw^* + 1$ according to their cumulative weights:

$$[1] \leftarrow n$$

$$(n(E), n(T), index) \leftarrow (1, 0, n - 1)$$

* currently, $n(E) = 1$ E-jobs, $n(T) = 0$ T-jobs

REPEAT

* compare the next cumulative weights $((n(E) + 1) - 1)\alpha$ and $(n(T) + 1)\beta$

IF $n(E)\alpha < (n(T) + 1)\beta$

THEN $n(E) \leftarrow n(E) + 1$, $[n(E)] \leftarrow index$

ELSE $n(T) \leftarrow n(T) + 1$, $[n - n(T) + 1] \leftarrow index$

$index \leftarrow index - 1$

UNTIL $index = nw^*$

3. Assign the position for W-jobs:

WHILE $index > 0$ DO

IF $(n(E))\alpha \leq (n(T) + 1)\beta$

THEN $[(n - n(T) + 1 - index)] \leftarrow index$, $index \leftarrow index - 1$

ELSE $[(n(E) + index)] \leftarrow index$, $index \leftarrow index - 1$

4. $C_{[1]} = p_{[1]}$

FOR $i = 2$ TO n DO

$$C_{[i]} = C_{[i-1]} + p_{[i]}$$

5. Determine the location of the due window:

IF $(n(E)\alpha \leq (n(T) + 1)\beta)$

THEN $d = (\sum_{\{j \leq n(E) + nw^*\}} p_{[j]}), e = d - K$

ELSE $e = (\sum_{\{j \leq n(E)\}} p_{[j]}), d = e + K$

Theorem 1 Algorithm 1 produces a schedule which minimizes $\sum_{j=1}^n (\alpha * E_j + \beta * T_j)$ in linear time ($O(n)$) provided the jobs are indexed in SPT order. If the jobs are not indexed in SPT order, an additional effort of $O(n * \log(n))$ for sorting is needed.

Proof. In the first step, the number nw^* is calculated according to its definition. According to Lemma 4, this is also the number of W -jobs. At all times, $n(E)$ is the number of E -jobs and $n(T)$ is the number of T -jobs. The E -jobs and T -jobs are sequenced according to their cumulative weights in accordance with Lemma 3.

The complexity of the algorithm is linear because each of the five steps in the algorithm is bounded by $O(n)$ operations. In the case job indices are not sorted in SPT order, we need an additional effort of order $O(n * \log(n))$. **QED**

Remark 2 It can be shown that the above algorithm constructs a schedule σ with $n(E(\sigma)) = \lceil (\beta(n - nw^*)) / (\alpha + \beta) \rceil$ E -jobs and $n(T(\sigma)) = n - \lceil (\beta(n - nw^*)) / (\alpha + \beta) \rceil - nw^*$ T -jobs where $\lceil x \rceil = \min\{k : k \geq x \text{ and integer}\}$. Furthermore, it can be shown that for any optimal schedule, $\min(n(E)\alpha, (n(T) + 1)\beta)$ is uniquely defined.

2.4 Parameter Case

Hall, Kubiak, and Sethi [56] used Dynamic Programming to solve the unit-weight case. Since the unit weight case is NP-complete ([56]), we present a dynamic programming algorithm. Our contribution is to extend the common due date to the common due window. We provide a dynamic programming algorithm of complexity $O(e * n^2 + K * n)$.

Remark 3 For the due-date parameter case, parameter d is non-restrictive if $d \geq MS$ holds. Hence, it can be treated like a decision variable case. Since this also holds for the due-window case, we assume here that $d < MS$ holds.

According to Lemma 1 (b), an optimal schedule exists in which either $C_j = d$ or $C_j = e$ for some j , or $S_{[1]} = 0$ holds. Recall that A_l is defined by $A_l = \sum_{j=1}^l p_j$.

Algorithm 2a: $C_j = d$ for some j

Considering the case $C_j = d$ for some j , we define the recursive function for the dynamic program as follows:

$g(l, t_{NT})$ is the minimal cost to schedule jobs 1, ..., l such that

$$t_{NT} = \sum_{(j \leq l \text{ and } C_j \leq d)} p_j$$

is the total processing time of non-tardy jobs.

Initial Conditions:

$$g(0, t_{NT}) = \begin{cases} 0 & \text{if } t_{NT} = 0 \\ \infty & \text{otherwise} \end{cases}$$

Recursive Relation:

$$g(l, t_{NT}) = \min \begin{cases} g(l-1, t_{NT} - p_l) + \alpha * (t_{NT} - p_l - K)^+ \\ g(l-1, t_{NT}) + \beta * (A_l - t_{NT}) \end{cases}$$

Answer:

$$Z_{2a} = \min_{(t_{NT}=K, \dots, e+K)} g(n, t_{NT})$$

Justification

Lemma 1(c)(ii) justifies job sequencing from the inside out.

Complexity

The complexity of the Dynamic Program is $O(n * d)$.

Algorithm 2b: $C_j = e$ for some j

For the case $C_j = e$ for some j , we define the recursive function for the dynamic program as follows:

$h(l, t_E)$ is the minimal cost to schedule jobs $1, \dots, l$ such that $t_E = \sum_{(j \leq l \text{ and } C_j \leq e)} p_j$
 is the total processing time of E -jobs.

Initial Conditions:

$$h(0, t_E) = \begin{cases} 0 & \text{if } t_E = 0 \\ \infty & \text{otherwise} \end{cases}$$

Recursive Relation:

$$h(l, t_E) = \min \begin{cases} h(l-1, t_E - p_l) + \alpha * (t_E - p_l) \\ h(l-1, t_E) + \beta * (A_l - t_E - K)^+ \end{cases}$$

Answer:

$$Z_{2b} = \min_{(t_E=0, \dots, e)} h(n, t_E)$$

Justification

Sequencing proceeds in SPT order from the inside out according to Lemma 1(c)(i).

Complexity

The complexity of the Dynamic Program is $O(n * e)$.

Algorithms for the Case $S_{[1]} = 0$

For this case, we distinguish whether a job crosses the due window, i.e., whether $S_k < e$ and $C_k > d$ for some job J_k , or not. If no job crosses the due window, then we must have $e \leq C_j \leq d$ for some j . For certain problem instances, only one of these cases may apply. For example, if $p_{max} < K$, we will not need Algorithm 2d. And if all processing times are multiples of 10 and $[e, d] = [24, 26]$, we will not need Algorithm 2c.

Algorithm 2c: $S_{[1]} = 0$ and $e \leq C_j \leq d$ for some j

In this case, we define function g as follows, where t_E refers to the cumulative processing time of jobs with $S_j < e$:

$g(l, t_E)$: is the minimal cost to schedule jobs $l, \dots, n$ such that $\sum_{(S_j < e, j \geq l)} p_j = t_E$

Initial Conditions:

$$g(n+1, t_E) = \begin{cases} 0 & \text{for } t_E = 0 \\ \infty & \text{otherwise} \end{cases}$$

Recursive Relation:

$$g(l, t_E) = \min \begin{cases} g(l+1, t_E - p_l) + \alpha * (e - t_E)^+ \\ g(l+1, t_E) + \beta * (A_l + t_E - d)^+ \end{cases}$$

Answer:

$$\min_{(t_E = e \dots d)} g(1, t_E).$$

Justification

Using Lemma 1, parts (c)(i) and (c)(ii), jobs with $S_j < e$ are sequenced in LPT order, the remaining ones in SPT order. The cumulative processing time of jobs with $S_j < e$, in the case where there is at least one on-time job, ranges from e to $\min(d, e + p_{max})$. We define the recursive function by scheduling jobs in LPT order from the outside toward the due window by backward dynamic programming.

Complexity Analysis

The complexity is $O(n * d)$.

Algorithm 2d: $S_{[1]} = 0$, $S_k < e$ and $C_k > d$ for some k

If no job completes on-time, there must be a job J_k which crosses the due window. For notational simplification in the dynamic program, we define $A_i^k = A_i + p_k$ for $i < k$ and $A_i^k = A_i$ for $i \geq k$. For all candidate jobs J_k with $p_k > d - e$, we use the following backward dynamic program:

$h_k(l, t_E)$: is the minimal cost to schedule jobs $l, \dots, n$ except job J_k such that

$$t_E = \sum_{(j \geq l, j \neq k, C_j < e)} p_j$$

Initial Conditions:

$$h_k(n+1, t_E) = \begin{cases} 0 & \text{for } t_E = 0 \\ \infty & \text{otherwise} \end{cases}$$

Recursive Relation:

$$h_k(l, t_E) = \min \begin{cases} h_k(l+1, t_E - p_l) + \alpha * (e - t_E) & l \neq k \\ h_k(l+1, t_E) + \beta * (A_l^k + t_E - d) & l \neq k \\ h_k(l+1, t_E) & l = k \end{cases}$$

Answer:

$$Z_{2d} = \min_k \min_{(t_E = d - p_k, \dots, e)} (h_k(1, t_E) + \beta * (t_E + p_k - d)).$$

Justification: The sequencing of E -jobs is according to Lemma 1(c)(i) and of the T -jobs according to Lemma 1(c)(ii). We add the cost for job J_k separately (see equation for Z_{2d}) in order to avoid an additional case distinction for h_k .

Complexity Analysis for Algorithm 2d

According to the state space, the complexity is $O(n * e)$ for each function h_k . Since the number of jobs for which $p_j > K$ holds is less or equal to n , the effort to evaluate Z_{2d} is bounded above by $O(n^2 * e)$.

The solution of the ordered weight case is as follows:

$$Z_2 = \min(Z_{2a}, Z_{2b}, Z_{2c}, Z_{2d}).$$

Complexity Analysis and Justification

The individual complexities are $O(n * d)$, $O(n * e)$, $O(n * d)$, and $O(n^2 * e)$, respectively, resulting in $O(n^2 * e + n * d)$.

Remark 4 If $K \geq p_{max}$, there is always a job which completes on time and Algorithm 2d is not needed. The complexity therefore reduces in this case ($K \geq p_{max}$) to $O(n * d)$ which is more efficient than in the common due-date case.

2.5 Final Remarks and Conclusions

Since Just-In-Time scheduling is a special case of Just-In-Window scheduling, we might expect those scheduling problems to be more complex. However, for the constant weight decision variable case, we are still able to provide a powerful $O(n)$ algorithm, provided the indices are sorted according to SPT. For the parameter case, a large due window ($K \geq p_{max}$) yields a more efficient algorithm than the corresponding problem for the common due date case.

The extension of this problem to parallel machines is treated in Krämer and Lee (1994), where we show that even the two machine decision variable case is NP-hard, and provide a dynamic program for the two machine case, as well as a heuristic with error bound for the general m machine case. Future directions include the development of dynamic programs and heuristics with error bounds for more general cost structures. The objective function for the decision variable case can also be extended to include a penalty for the assignment of the location of the due window.

CHAPTER 3

PARALLEL DUE WINDOW SCHEDULING

3.1 Introduction

In this chapter, we are considering one aspect of World-Class Manufacturing. Its pillars are Total Quality Control (TQC), Just-In-Time (JIT) Production, Factory Automation (FA), and Total Productive Maintenance (TPM). One goal of Just-In-Time Production is to eliminate all inventory. Inventory often binds significant portions of capital and for many companies, reduction of inventory is an existential problem. Other problems are also addressed by the Just-In-Time concept. These include rework, machine downtime, change order, long setups, vendor delinquencies, etc. Just-In-Time Production Scheduling is modeled by nonregular performance measures, where both earliness and tardiness penalties are included. Many publications have dealt with various earliness and tardiness problems. Baker and Scudder give an excellent review in [20]. For the single machine case, the first pseudopolynomial algorithms were presented by Hall and Posner [55] and Hall, Kubiak, and Sethi [56]. Lee, Danusaputro, and Lin [77] provide a pseudopolynomial dynamic program for a more general objective function. A heuristic with finite error bound is presented in Liman and Lee [80]. Concerning monotone penalty functions with respect to earliness and tardiness, Federgruen and Mosheiov [39] present optimality properties.

This problem is extended in two ways: First, the due date formulation is generalized to include the notion of due windows where the term due window refers to a time interval during which no penalty is incurred. Second, the results are transferred from single machines to parallel machines.

Concerning due window scheduling, Anger, Lee, and Martin-Vega [6] consider a problem where the objective is to minimize the number of jobs completing outside their due windows,

where the largest window size is less than or equal to the smallest processing time. Sidney [102] and Lakshminarayan et al. [74] considered another minimax formulation where the objective is to minimize the maximum earliness and tardiness penalty. Restrictions are target starting time and target due date. The idea of due window concepts was implicitly used. Cheng [26] studies the common due window problem with the constraint that the due window is smaller than the smallest processing time. Unit weights with respect to the middle of the window are used in the objective function. Lee [76] has modeled and solved several problems in the area of constant window scheduling. Krämer and Lee [72] studied the single machine common due window problem for constant weights. A polynomial algorithm has been provided for a problem with constant weights.

The second extension is from single machines to parallel machines. Emmons [38] has dealt with the problem of constant earliness and tardiness weights with **common due date**. As for the one machine problem, the problem is solved by assigning the jobs in non-increasing processing time order to the sequence of non-decreasing **cumulative weights**. In addition, Emmons applied secondary criteria since many optimal solutions exist because the cumulative weights occur in multiples of m .

While the extension from due date to due window preserves the polynomial complexity in the single machine case for constant weights, this is no longer true in general for the parallel machine case. For the constant weight decision variable case with a common due date, the parallel machine problem is polynomial [38]. But the extension to the due window case and parallel machines is NP-hard. This holds even for two machines and unit weights.

This chapter is organized as follows: In Section 3.2, we introduce notation and review general properties from the single machine case since they also apply to the parallel machine case. In Section 3.3, we deal with the complexity of the problem and develop optimality properties for parallel machines. For the two machine case, a dynamic program and a heuristic is presented in Section 3.4 and the extension of the heuristic to an arbitrary number of parallel machines is given in Section 3.5, including an analysis of the absolute

and relative error. We also show that the relative error vanishes as the number of jobs increases.

3.2 Notation and General Properties

We assume that all data used in this chapter are integral. We have n jobs, and for each job J_j ($j = 1, \dots, n$), we use the following job attributes:

1. Processing time p_j
2. Starting time S_j
3. Completion time C_j

We also define $p_{max} = \max\{p_j : 1 \leq j \leq n\}$.

For each job, α is the unit earliness penalty and β is the unit tardiness penalty, i.e., we consider the constant weight case in this chapter. A window is defined as an interval $[\epsilon, d]$, where ϵ is the earliest due date and d is the latest due date. The size of the due window $K = d - \epsilon$ is always considered as a given parameter of the problem. However, the location of the due window is considered as a decision variable, i.e., $d(\epsilon)$ is a decision variable. Jobs completing within the due window do not incur a penalty. The earliness and tardiness of a job is defined as $E_j = (\epsilon - C_j)^+$ and $T_j = (C_j - d)^+$, respectively, where $x^+ = \max(x, 0)$. The penalty of job J_j is a function of the completion time C_j and is defined as $\alpha(\epsilon - C_j)^+ + \beta(C_j - d)^+ = \alpha E_j + \beta T_j$. The number of parallel machines is denoted by m . The i^{th} machine is denoted by M^i .

We are given the size of the due window $d - \epsilon$. The due window is common to all jobs. We assume that no costs are incurred in assigning the location of the due window. Our purpose is to determine the location of the due window and to find a schedule to minimize

$$Z = \sum_{j=1}^n (\alpha(\epsilon - C_j)^+ + \beta(C_j - d)^+) = \sum_{j=1}^n (\alpha E_j + \beta T_j).$$

Throughout the chapter, we assume without loss of generality that the jobs are indexed in shortest processing time order (SPT), i.e., $p_j \leq p_{j+1}$ for all j . We denote the processing

time of the first k jobs by A^k , i.e., $A^k = \sum_{j=1}^k p_j$. For a given schedule, $[j]^i$ denotes the j^{th} job on machine M^i in that schedule.

For the due window problem, we have the following properties for each machine [72]:

Property 1: There exists an optimal schedule without machine idle time between the first job and the last job.

Property 2: In any optimal schedule, jobs completing before e ($C_j \leq e$) are sequenced in *reverse SPT* order, and jobs which start after d ($S_j \geq d$) are sequenced in SPT order.

Property 3: There exists an optimal schedule where one job finishes at either e or d :

- either (i) $C_j = d$ for some j
 or (ii) $C_j = d - K = e$ for some j

Remark 5 In order to find an optimal schedule, we can restrict ourselves to schedules which satisfy the Properties 1-3. Hence, throughout the entire chapter, we consider only schedules where for each machine M^i , there exists some job J_j such that $C_j = e$ or $C_j = d$ holds.

The main difference between common due date and common due window scheduling is the determination of the set of jobs W completing within the due window. For a schedule σ , we define the following sets of jobs:

$$E(\sigma) = \{J_j : C_j(\sigma) \leq e\}$$

$$T(\sigma) = \{J_j : S_j(\sigma) \geq d\}$$

$$W(\sigma) = \{J_j : C_j(\sigma) > e \text{ and } S_j(\sigma) < d\}$$

$$WI(\sigma) = \{J_j : S_j(\sigma) \geq e \text{ and } C_j(\sigma) \leq d\}$$

Jobs in E are completed at or before the window and jobs in T start at or after the window. Set W contains the remaining jobs. Set WI contains those jobs which are completely processed within the due window. According to the job sets, we refer to the jobs as *E-jobs*, *T-jobs*, *W-jobs*, or *WI-jobs*. Roughly speaking, *E-jobs* are early and *T-jobs* are tardy.

Strictly speaking, the last *E-job* may be completed at the beginning of the due window e and the last *W-job* may be tardy. For machine M^i , we define $JS^i(\sigma)$ as follows:

$$JS^i(\sigma) = \{J_j : J_j \text{ is scheduled on } M^i\}$$

Accordingly, we define $E^i(\sigma) = JS^i(\sigma) \cap E(\sigma)$, $T^i(\sigma) = JS^i(\sigma) \cap T(\sigma)$, $W^i(\sigma) = JS^i(\sigma) \cap W(\sigma)$, and $WI^i(\sigma) = JS^i(\sigma) \cap W(\sigma)$.

When there is no ambiguity concerning the schedule, we simply write JS^i , E^i , T^i , W^i and WI^i . Furthermore, denote the cardinality of a set S by $n(S)$, i.e., $n(S) = |S|$. By the above definitions, we have $n(JS^i) = n(E^i) + n(T^i) + n(W^i)$.

In order to express the objective function in terms of processing times (compare [72]), and to simplify the discussion, we define cumulative weights. For the definition, please recall that we restrict ourselves to schedules which satisfy *Properties 1-3*. According to *Property 3*, such schedules satisfy either $C_j = e$ or $C_j = d$ for some job J_j .

Definition 4 For a schedule σ , we define the cumulative weight $cw(J_j, \sigma)$ for job J_j as follows:

1. If job J_j is the k^{th} E-job on machine M^i in schedule σ , i.e., $J_j \in E^i(\sigma)$ and $[k]^i = j$, we define

$$cw(J_j, \sigma) = (k - 1)\alpha.$$

2. If job J_j is the k^{th} last T-job on machine M^i in schedule σ , i.e., $J_j \in T^i(\sigma)$ and $[n(M^i) + 1 - k]^i = j$, we define

$$cw(J_j, \sigma) = k\beta.$$

3. If job J_j is a W-job on machine M^i in schedule σ , i.e., $J_j \in W^i(\sigma)$, we define

$$cw(J_j, \sigma) = \begin{cases} (n(E^i(\sigma)))\alpha & \text{if } C_k = d \text{ for some } J_k \text{ on machine } M^i \\ (n(T^i(\sigma)) + 1)\beta & \text{otherwise} \end{cases}$$

We also define

$$cw(W^i, \sigma) = \begin{cases} (n(E^i(\sigma)))\alpha & \text{if } C_k = d \text{ for some } J_k \text{ on machine } M^i \\ (n(T^i(\sigma)) + 1)\beta & \text{otherwise} \end{cases} \quad (3.1)$$

since the cumulative weight of the W-jobs is the same for every W-job in W^i .

3.3 Complexity and Optimality Properties

We first demonstrate that the problem for the 2-machine case is NP-hard. As is shown by Krämer and Lee [72], the constant weight decision variable case for a single machine can be extended from the due date case to the due window case while it remains polynomially solvable. However, such a result cannot be obtained for the parallel machines case. Even for the unit weight case where $\alpha = \beta = 1$, the problem is NP-hard:

Theorem 2 *The minimization of $\sum_j (E_j + T_j)$ on two parallel machines is NP-hard.*

Proof. We reduce the Partition problem to our problem. An instance of the Partition problem is given by $a_1, \dots, a_N$ and by the question: Does there exist a subset $I \subset \{1, \dots, N\}$ such that $\sum_{j \in I} a_j = A/2$ where $A = \sum_{j=1}^N a_j$?

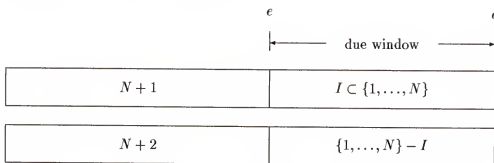


Figure 3.1. Problem complexity for parallel machines

Let the window size be $d - \epsilon = A/2$, the number of jobs to be scheduled $n = N + 2$ with $p_j = a_j$ for $j = 1, \dots, N$, and $p_{N+1} = p_{N+2} = A/2 + 1$ be an instance of the scheduling problem. Figure 3.1 illustrates the construction of the scheduling instance. Clearly, a zero penalty solution for the scheduling problem exists if and only if a partition exists. **QED**

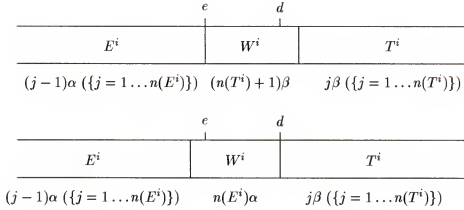


Figure 3.2. Cumulative weights for two solution structures

Please note that if m , the number of machines, is not fixed, the problem is NP-hard in the strong sense. This can be shown by reducing the problem **3-Partition** to this problem. In the next lemma, we rewrite the objective function in terms of cumulative weights.

Lemma 5 *Let σ be a schedule which satisfies Properties 1-3. The weighted earliness and tardiness penalty can be expressed by cumulative weights for the processing times as follows:*

$$\sum_{j=1}^n (\alpha E_j(\sigma) + \beta T_j(\sigma)) = \sum_{i=1}^m \left\{ \left(\sum_{J_j \in JS^i(\sigma)} cw(J_j, \sigma) * p_j \right) - cw(W^i, \sigma) * (d - e) \right\} \quad (3.2)$$

Proof. See [72] and apply the result for each machine.

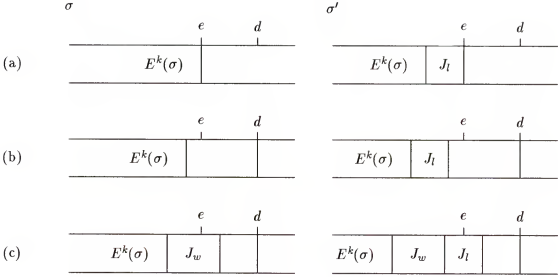
QED

Remark 6 Figure 3.2 illustrates the cumulative weights for the two situations $C_j = e$ and $C_j = d$ for some job J_j . Note that the W -jobs contribute

$$\min(\alpha(n(E^i)), \beta(n(T^i) + 1)) \left(\sum_{J_j \in W^i} p_j - (d - e) \right)$$

toward the objective function. This explains why W -jobs also have a cumulative weight.

In the next lemma, we prove that the number of E -jobs on any two machines differs by no more than one. This is true for all optimal schedules. The same statement also holds for T -jobs.

Figure 3.3. Machine M^k for schedule σ and σ'

Lemma 6 For every optimal schedule, we have

$$\max_{M^i} n(E^i(\sigma)) - \min_{M^i} n(E^i(\sigma)) \leq 1 \quad (3.3)$$

$$\max_{M^i} n(T^i(\sigma)) - \min_{M^i} n(T^i(\sigma)) \leq 1 \quad (3.4)$$

Proof. Assume that there is an optimal schedule σ which satisfies $\max_{(M^i)} n(E^i(\sigma)) - \min_{(M^i)} n(E^i(\sigma)) > 1$. Let M^i and M^k be machines with $n(E^i(\sigma)) = \max_{(M^i)} n(E^i(\sigma))$ and $n(E^k(\sigma)) = \min_{(M^i)} n(E^i(\sigma))$. Let job J_l be the E -job on machine M^i with $cw(J_l, \sigma) = (n(E^i(\sigma)) - 1)\alpha$. Obtain schedule σ' from schedule σ by removing job J_l from machine M^i and delaying the other E -jobs of machine M^i by p_l time units.

We can disregard jobs in $JS^i(\sigma')$. Their contribution toward the objective function is either the same or less, since (1) the E -jobs and T -jobs of machine M^i have the same cumulative weights as before, and (2) the W -jobs of machine M^i satisfy $cw(W^i, \sigma') - cw(W^i, \sigma) \in \{0, -\alpha\}$ (0, if $C_j(\sigma) = \epsilon$ for some job $J_j \in E^i(\sigma)$ and $-\alpha$ otherwise).

Job J_l will be processed on machine M^k instead as described below. Since other jobs have not been affected, we can therefore concentrate on the jobs $JS^k(\sigma')$. We distinguish the following two cases:

1. $C_j = e$ for some job J_j on machine M^k .

In this case, we schedule job J_l on machine M^k such that $C_l(\sigma') = e$ (see Figure 3.3, part (a)). Notice that the only job on machine M^k for which the cumulative weight has changed is job J_l : $cw(J_l, \sigma') - cw(J_l, \sigma) = (n(E^k(\sigma)) - (n(E^i(\sigma)) - 1))\alpha < 0$. Therefore, we have

$$cost(\sigma') - cost(\sigma) \leq (cw(J_l, \sigma') - cw(J_l, \sigma))p_l < 0.$$

2. $C_j \neq e$ for all jobs J_j on machine M^k .

According to Property 3, $C_j \neq e$ for all jobs J_j on machine M^k implies that we have $C_j = d$ for some job J_j on this machine. In this case, the cumulative weights of the W -job for machine M^k may change. We distinguish the following cases:

- (a) $\sum_{J_j \in W^k(\sigma)} p_j - (d - e) < p_l$:

Schedule job J_l on machine M^k such that $C_l(\sigma') = \min_{J_j \in W(\sigma')} S_j(\sigma)$ (see Figure 3.3, part (b)). The jobs on machine M^k with modified cumulative weight are J_l and the set $W^k(\sigma)$. Since $cw(J_l, \sigma) = (n(E^i(\sigma)) - 1)\alpha$, we have

$$cw(W^k, \sigma') - cw(W^k, \sigma) = \alpha.$$

$$cw(J_l, \sigma') - cw(J_l, \sigma) = (n(E^k(\sigma))\alpha - (n(E^i(\sigma)) - 1)\alpha) \leq -\alpha.$$

Therefore, we have

$$\begin{aligned} cost(\sigma') - cost(\sigma) &\leq p_l(-\alpha) + \left(\sum_{j \in W^k(\sigma)} p_j - (d - e) \right) \alpha \\ &= (p_l - \left(\sum_{j \in W^k(\sigma)} p_j - (d - e) \right))(-\alpha) < 0 \end{aligned}$$

- (b) $\sum_{J_j \in W^k(\sigma)} p_j - (d - e) \geq p_l$:

Let J_w be the job which starts before the due window in schedule σ . Since we have $C_w \neq e$, we have $\sum_{J_j \in W(\sigma)} p_j - p_w < d - e$. In this case, obtain schedule σ' by scheduling job J_l such that $C_l(\sigma') = C_w(\sigma)$ according to Figure 3.3, part (c).

The cumulative weights of the former E -jobs and T -jobs remain unaltered. For

job J_w , we have also $cw(J_w, \sigma') = cw(J_w, \sigma)$, since the cumulative weight of job J_w depends only on the number of E -jobs preceeding it (compare Figure 3.2). And this number is the same in both schedules. We have two further cases to distinguish:

- i. $n(E^k(\sigma')) = n(E^k(\sigma)) + 1$:

In this case, job J_w is an E -job in schedule σ' as illustrated in Figure 3.3, part (c). For the W -jobs of machine M^k w.r.t. schedule σ' , we have $cw(W, \sigma') = (n(E^k(\sigma)) + 1)\alpha$. Further, $cw(W, \sigma') - cw(W, \sigma) = \alpha > 0$, $\sum_{J_j \in W^k(\sigma) - \{J_w\}} p_j < (d - e)$, and $cw(J_l, \sigma') \leq cw(J_l, \sigma)$ imply

$$\begin{aligned}
 & cost(\sigma') - cost(\sigma) \\
 & \leq \{ \sum_{(J_j \in W^k(\sigma'))} (cw(J_j, \sigma') - cw(J_j, \sigma)) p_j \} \\
 & \quad - (cw(W, \sigma') - cw(W, \sigma))(d - e) \\
 & = \alpha \{ \sum_{(J_j \in W^k(\sigma) - \{J_w\})} p_j - (d - e) \} \\
 & \quad + (cw(J_l, \sigma') - cw(J_l, \sigma)) p_l \\
 & < 0.
 \end{aligned}$$

- ii. $n(E^k(\sigma')) = n(E^k(\sigma))$:

That is, the number of E -jobs on machine M^k is unchanged. It implies that job J_w remains to be a W -job and job J_l starts after e . This case is not represented by Figure 3.3, part (c). Except job J_l , all jobs in machine M^k have the same cumulative weight. Since $cw(J_l, \sigma) = (n(E^i(\sigma)) - 1)\alpha$, we have

$$\begin{aligned}
 & cost(\sigma') - cost(\sigma) \\
 & = (cw(J_l, \sigma') - cw(J_l, \sigma)) p_l \\
 & = (n(E^k(\sigma)) - (n(E^i(\sigma)) - 1)) \alpha p_l < 0.
 \end{aligned}$$

Thus, in all cases, we can find a better schedule, which contradicts the assumption that σ is optimal. Therefore, we must have $\max_{M^i} n(E^i(\sigma)) - \min_{M^i} n(E^i(\sigma)) \leq 1$. The proof for $\max_{M^i} n(T^i(\sigma)) - \min_{M^i} n(T^i(\sigma)) \leq 1$ is symmetric to the proof for $\max_{M^i} n(E^i(\sigma)) - \min_{M^i} n(E^i(\sigma)) \leq 1$. **QED**

Remark 7 In order to characterize the W-jobs, we apply Lemma 4 in [72] for every machine and have immediately that there is an optimal schedule σ which satisfies

$$p_j \leq p_l \quad \text{for all } J_j \in W^i(\sigma) \quad \text{and for all } J_l \in E^i(\sigma) \cup T^i(\sigma) \quad (3.5)$$

for all machines M^i . The following Lemma will extend this statement.

Lemma 7 There exists an optimal schedule σ^* which satisfies

$$p_j \leq p_l \quad \text{for all } J_j \in W(\sigma^*) \quad \text{and for all } J_l \in E(\sigma^*) \cup T(\sigma^*). \quad (3.6)$$

Proof. To facilitate the proof, we introduce the following definition: For a given schedule σ , define the set $WET(\sigma)$ as follows:

$$WET(\sigma) = \{(J_j, J_l) : p_j > p_l, \quad J_j \in W(\sigma) \quad \text{and} \quad J_l \in E(\sigma) \cup T(\sigma)\} \quad (3.7)$$

In other words, $WET(\sigma)$ is the set of job-pairs which violate Equation (3.6). Therefore, we need to prove that there exists an optimal schedule σ^* for which $n(WET(\sigma^*)) = 0$ holds where $n(WET(\sigma^*))$ denotes the number of pairs in $WET(\sigma^*)$. Let $\sigma^{\min}$ be an optimal schedule which satisfies

$$n(WET(\sigma^{\min})) = \min\{n(WET(\sigma)) : \sigma \text{ is an optimal schedule}\}.$$

If $n(WET(\sigma^{\min})) = 0$, we are done, i.e., there exists an optimal schedule which satisfies (3.7).

Now assume that $n(WET(\sigma^{\min})) > 0$ and that $(J_j, J_l) \in WET(\sigma^*)$, i.e., $J_j \in W^i(\sigma^*)$ and $J_l \in E^k(\sigma^*) \cup T^k(\sigma^*)$ for some machines M^i and M^k . We show a contradiction by constructing an optimal schedule σ^+ which satisfies $n(WET(\sigma^+)) < n(WET(\sigma^{\min}))$, i.e.,

which has a fewer number of pairs which violate Equation (3.7). According to the proof of Lemma 4 in [72], we can restrict ourselves to the case where schedule σ^{min} satisfies Equation (3.5) for each machine. In the following, we write W^i , E^k , etc. for the sets which refer to schedule σ^{min} . We first distinguish between $J_l \in E^k$ and $J_l \in T^k$:

1. There exist two jobs $J_j \in W^i$ and $J_l \in E^k$ with $p_j > p_l$:

We distinguish the following cases:

- (a) $cw(J_l, \sigma^{min}) \leq \alpha(n(E^i) - 1)$:

Obtain schedule σ' by interchanging job J_l with the E -job on machine M^i which has the same cumulative weight. By construction, we have $cost(\sigma') = cost(\sigma^{min})$. Note that schedule σ' violates Equation (3.5) for machine M^i . Now observe that by the same interchange argument as in the proof of Lemma 4 in [72], we can construct a schedule σ^+ with $cost(\sigma^+) \leq cost(\sigma')$ and $n(WET(\sigma^+)) < n(WET(\sigma^{min}))$. This holds because there are no job-pairs which violate Equation (3.6) in schedule σ^+ , but not in schedule σ^{min} , i.e., we have $WET(\sigma^+) \subseteq WET(\sigma^{min})$. And there is at least one job pair $((J_j, J_l))$ - may be more - which violates Equation (3.6) in schedule σ^{min} , but not in schedule σ^+ . That is, σ^+ is an optimal schedule with fewer job pairs violating Equation (3.6) (since at least the pair (J_j, J_l) is removed from the set $WET(\sigma^{min})$). This establishes the desired contradiction.

- (b) $cw(J_l, \sigma^{min}) > \alpha(n(E^i) - 1)$:

In this case, we further distinguish as follows:

- i. $C_j = e$ for some job J_j on machine M^i :

Obtain schedule σ' by removing job J_l from machine M^k and placing it on machine M^i as E -job with cumulative weight $\alpha((n(E^i) - 1) + 1) (\leq cw(J_l, \sigma^{min}))$. By construction, we have $cost(\sigma') \leq cost(\sigma^{min})$. Yet schedule σ' violates Equation (3.5) for machine M^i . Therefore, we can construct

an optimal schedule σ^+ which satisfies $n(WET(\sigma^+)) < n(WET(\sigma^{min}))$, a contradiction.

- ii. $C_j \neq e$ for all jobs J_j on machine M^i :

For this case, we further distinguish as follows:

- A. $(\sum_{J_j \in W^i} p_j) - p_j + p_l < d - e$:

In this case, we construct schedule σ^+ by placing job J_l as W -job in machine M^i without changing the structure of machine M^i . Since

$$cw(J_l, \sigma^{min}) \geq (\alpha((n(E^i) - 1) + 1) = cw(J_l, \sigma^+),$$

the cumulative weight of job J_l is either maintained or has decreased.

The jobs on machine M^k have either the same or less cumulative weight.

Therefore, we have a schedule σ^+ with

$$cost(\sigma^+) \leq cost(\sigma') \quad \text{and} \quad n(WET(\sigma^+)) < n(WET(\sigma^{min})).$$

That is, σ^+ is an optimal schedule with fewer job pairs violating Equation (3.6).

- B. $(\sum_{J_j \in W^i} p_j) - p_j + p_l \geq d - e$:

In this case, we finally apply Lemma 6 and conclude by $n(E^k) > n(E^i)$ that $n(E^k) = n(E^i) + 1$. Therefore, we have $cw(J_j, \sigma^{min}) = \alpha(n(E^k) - 1) = \alpha(n(E^i)) = cw(J_l, \sigma^{min})$. In this case, we can construct schedule σ^+ by exchanging jobs J_l and J_j without changing the number of W -jobs, E -jobs or T -jobs on any machine. Since both jobs have the same cumulative weight in both schedules, we have $cost(\sigma^+) = cost(\sigma^{min})$. Since further $n(WET(\sigma^+)) < n(WET(\sigma^{min}))$, we have a contradiction to our assumption.

2. There exist two jobs $J_j \in W^i$ and $J_l \in T^k$ with $p_j > p_l$:

This proof for this case is symmetric to the proof above.

QED

In order to characterize the number of W -jobs in an optimal schedule, we introduce the quantity nw .

Definition 5 Let nw be defined as the maximum possible number of jobs which can be processed in time $m * (d - \epsilon)$, i.e.,

$$nw = \begin{cases} \max\{l : \sum_{j=1}^l p_j \leq m * (d - \epsilon)\} & \text{if } p_1 \leq m * (d - \epsilon) \\ 0 & \text{otherwise} \end{cases} \quad (3.8)$$

Lemma 8 Let σ be any schedule which satisfies Properties 1-3 and Equation (3.6). Then we have

$$n(W(\sigma)) = nw \quad \text{if and only if} \quad n(WI(\sigma)) = nw \text{ and } A^{nw} = m * (d - \epsilon). \quad (3.9)$$

Proof. By definition of W and nw , we have

$$\sum_{J_j \in W(\sigma)} p_j = \sum_{i=1}^m \sum_{J_j \in W^i(\sigma)} p_j \geq m * (d - \epsilon) \geq \sum_{j=1}^{nw} p_j. \quad (3.10)$$

Since schedule σ satisfies Equation (3.6), the set $W(\sigma)$ consists of the $n(W(\sigma))$ smallest jobs. Therefore, we have $n(W(\sigma)) \geq nw$.

$\rightarrow$: First note that by definition of W^i and by Property 1, we have $\sum_{J_j \in W^i(\sigma)} p_j \geq d - \epsilon$ for all i . If $n(W(\sigma)) = nw$, then the relations in (3.10) are equalities and therefore, $\sum_{J_j \in W(\sigma)} p_j = m * (d - \epsilon)$ holds. This implies $\sum_{J_j \in W^i(\sigma)} p_j = d - \epsilon$ and $WI^i(\sigma) = W^i(\sigma)$ for all i . Therefore, $n(WI(\sigma)) = n(W(\sigma)) = nw$ and $A^{nw} = m * (d - \epsilon)$ holds.

$\leftarrow$: Now note that by definition of WI^i , we have $\sum_{J_j \in WI^i(\sigma)} p_j \leq d - \epsilon$ for all i . If $n(WI(\sigma)) = nw$ and $A^{nw} = m * (d - \epsilon)$ holds, we have $\sum_{J_j \in WI(\sigma)} p_j = m * (d - \epsilon)$. This implies $\sum_{J_j \in WI^i(\sigma)} p_j = d - \epsilon$ and $W^i(\sigma) = WI^i(\sigma)$ for all i . Therefore, $n(W(\sigma)) = n(WI(\sigma)) = nw$ holds. QED

Lemma 9 Let σ be any schedule which satisfies Properties 1-3 and Equation (3.6). Then we have

$$nw - m < n(WI(\sigma)) \leq nw. \quad (3.11)$$

Proof. By definition of nw , we have $n(WI) \leq nw$ for every schedule and thus $n(WI(\sigma)) \leq nw$ holds.

1. $n(W(\sigma)) = nw$: If $n(W(\sigma)) = nw$ holds, we have $n(WI(\sigma)) = nw$ by Lemma 8.

Thus, Equation (3.11) is satisfied.

2. $n(W(\sigma)) > nw$: Since schedule σ satisfies *Properties 1* and *3*, we have $n(W^i(\sigma)) - n(WI^i(\sigma)) \leq 1$ for all i . Therefore, we have $n(W(\sigma)) \leq n(WI(\sigma)) + m$. This implies $nw < n(WI(\sigma)) + m$. Hence, Equation (3.11) holds as well.

QED

3.4 The Two Machine Case

In this section, we consider the two machine case. Recall that the single machine case is treated in polynomial time. Kanet [63] provided an $O(n * \log(n))$ algorithm for due date case, and Krämer and Lee [72] provided an $O(n * \log(n))$ algorithm for the due window case.

We use dynamic programming to determine the allocation of *WI-jobs* to the machines. While Lemma 9 gives a range for the number of *WI-jobs* for a general number of machines, the following lemma refers to the potential *WI-sets* which are of interest for the **Two Parallel Machine** case.

Lemma 10 *There exists an optimal schedule σ^* for the Two Parallel Machine problem which satisfies Properties 1-3, Equation (3.6), and the following statements concerning the set $WI(\sigma^*)$:*

1. *If $n(WI(\sigma^*)) = nw - 1$, then $n(W(\sigma^*)) = nw + 1$ and $WI(\sigma^*) \subset \{J_1, \dots, J_{nw}\}$.*
2. *If $n(WI(\sigma^*)) = nw$, then $nw \leq n(W(\sigma^*)) \leq nw + 2$ and $WI(\sigma^*) \subset \{J_1, \dots, J_{nw+1}\}$.*

Concerning the E-jobs and T-jobs, we have in case of $n(W(\sigma^*)) > nw$, that either $E(\sigma^*) \cup T(\sigma^*) = \{J_{nw+2}, \dots, J_n\}$ or $E(\sigma^*) \cup T(\sigma^*) = \{J_{nw+3}, \dots, J_n\}$ holds.

Proof. Consider a schedule σ^* which satisfies *Properties 1-3* and Equation (3.6).

According to Equation (3.11) in Lemma 9, we have $nw - 2 < n(WI(\sigma^*)) \leq nw$. Accordingly, we distinguish the cases $n(WI(\sigma^*)) = nw - 1$ and $n(WI(\sigma^*)) = nw$.

1. $n(WI(\sigma^*)) = nw - 1$:

In this case, we have $n(W(\sigma^*)) \leq nw + 1$ by *Property 3*. and $n(W(\sigma^*)) \neq nw$ by (3.9).

Hence, we have $n(W(\sigma^*)) = nw + 1$ and by Equation (3.6) $W(\sigma^*) = \{J_1, \dots, J_{nw+1}\}$.

Hence, there are two jobs which are *W-jobs*, but not *WI-jobs*. By interchange argument we can determine one of them: Job J_{nw+1} can always be scheduled as non-*WI-job* without changing the cost and without changing the number of *WI-jobs* or *W-jobs*.

Therefore, there exists an optimal schedule σ' which satisfies *Properties 1-3*, Equation (3.6), $n(W(\sigma')) = nw + 1$, $n(WI(\sigma')) = nw - 1$, and $WI(\sigma') \subset \{J_1, \dots, J_{nw}\}$.

2. $n(WI(\sigma^*)) = nw$:

(a) $A^{nw} = 2 * (d - \epsilon)$:

By (3.9), we have $n(W(\sigma^*)) = nw$, $WI(\sigma^*) = \{J_1, \dots, J_{nw}\}$, and

$$\{J_1, \dots, J_{nw}\} \subset \{J_1, \dots, J_{nw+1}\}.$$

(b) $A^{nw} < 2 * (d - \epsilon)$:

Since $\sum_{J_j \in W(\sigma^*)} p_j \geq 2 * (d - \epsilon)$ holds and since schedule σ^* satisfies Equation (3.6), we have $n(W(\sigma^*)) \geq nw + 1$. Since schedule σ^* satisfies *Properties 1-3*, we have by the definitions of W and WI that $n(W(\sigma^*)) \leq nw + 2$ holds. Therefore, we have $n(W(\sigma^*)) \in \{nw + 1, nw + 2\}$.

i. In case of $n(W(\sigma^*)) = nw + 1$, $WI(\sigma^*) \subset \{J_1, \dots, J_{nw+1}\}$ holds.

ii. In case of $n(W(\sigma^*)) = nw + 2$, there are two jobs which are *W-jobs*, but which are not *WI-jobs*. By interchange argument we can determine one of them: Job J_{nw+2} can always be scheduled as non-*WI-job* without changing the cost. Therefore, in case of $nw + 2$ *W-jobs*, there exists an optimal schedule σ' which satisfies *Properties 1-3*, Equation (3.6), and $WI(\sigma') \subset \{J_1, \dots, J_{nw+1}\}$.

Now consider the *E-jobs* and *T-jobs*. If $n(W(\sigma^*)) > nw$, we have that either $n(W(\sigma^*)) = nw + 1$ or $n(W(\sigma^*)) = nw + 2$ holds. If $n(W(\sigma^*)) = nw + 1$, we have by Equation (3.6) that $E(\sigma^*) \cup T(\sigma^*) = \{J_{nw+2}, \dots, J_n\}$ holds. And if $n(W(\sigma^*)) = nw + 2$, we have by Equation (3.6) that $E(\sigma^*) \cup T(\sigma^*) = \{J_{nw+3}, \dots, J_n\}$ holds. **QED**

Dynamic Program

According to the previous lemma, we must examine the cases

1. $n(WI) = nw - 1$ and $WI \subset \{J_1, \dots, J_{nw}\}$.

In this case, we have nw different sets for the *WI-jobs*. We obtain those sets by selecting the job J_k which is not a *WI-job*. Then, we have $WI = \{J_1, \dots, J_{nw}\} - \{J_k\}$.

This is done for each $J_k \in \{J_1, \dots, J_{nw}\}$.

2. $n(WI) = nw$ and $WI \subset \{J_1, \dots, J_{nw+1}\}$.

In the same manner, we select a job $J_k \in \{J_1, \dots, J_{nw+1}\}$ which is not a *WI-job*.

Hence, the remaining jobs $\{J_1, \dots, J_{nw+1}\} - \{J_k\}$ are the *WI-jobs*. Again, this is done for every job $J_k \in \{J_1, \dots, J_{nw+1}\}$.

The following dynamic program is used to determine whether the potential *WI-sets* can indeed be scheduled within the window. The function $f_k(l, t)$ depends on job J_k , index l ($\leq nw + 1$), and parameter t . The function determines whether jobs $J_1, \dots, J_l$, except job J_k ($k \in \{1, \dots, nw + 1\}$) can be scheduled as *WI-jobs* such that the cumulative processing time for machine M^1 equals a given parameter t , i.e., $\sum_{J_j \in WI^1} p_j = t$.

Algorithm 1: Feasibility of *WI-jobs*

$f_k(l, t) = YES$ if there is a subset of jobs $S \subset \{J_1, \dots, J_l\} - \{J_k\}$ such that $\sum_{j \in S} p_j = t$
and $\sum_{j=1, j \neq k}^l p_j - t \leq d - e$ and $f_k(l, t) = NO$ otherwise.

The jobs in S correspond to the ones which are scheduled in the due window of machine M^1 .

In determining the optimal schedule for the 2 machine problem, we are interested in the

function values $f_k(nw, t)$ for all $k \in \{1, \dots, nw\}$ and all $t \in \{(A^{nw} - p_k - (d - e))^+, \dots, d - e\}$ and $f_k(nw + 1, t)$ for all $k \in \{1, \dots, nw + 1\}$ and all $t \in \{(A^{nw+1} - p_k - (d - e))^+, \dots, d - e\}$.

These values of t will guarantee that a *YES* answer can only occur when $\sum_{J_j \in WI} p_j \leq d - e$ holds for both machines $i \in \{1, 2\}$. Note that these values of t cover the range of the total processing times of *WI-jobs* in machine M^1 .

Initial Conditions:

$$f_k(0, t) = \begin{cases} YES & t = 0 \\ NO & \text{otherwise} \end{cases}$$

Recursive Relation:

$$f_k(l, t) = \begin{cases} f_k(l - 1, t) & \text{if } l = k \\ YES & \text{if } l \neq k \text{ and } f_k(l - 1, t - p_l) = YES \\ YES & \text{if } l \neq k \text{ and } f_k(l - 1, t) = YES \text{ and } \sum_{j=1, j \neq k}^l p_j - t \leq d - e \\ NO & \text{otherwise} \end{cases}$$

where $l = 0, \dots, nw + 1$ and $t = 0, \dots, d - e$.

Answer:

The results of this dynamic program are used in various places. As mentioned above, we are interested in $f_k(nw, t)$ for all $k \in \{1, \dots, nw\}$ and all $t \in \{(A^{nw} - p_k - (d - e))^+, \dots, d - e\}$ and $f_k(nw + 1, t)$ for all $k \in \{1, \dots, nw + 1\}$ and all $t \in \{(A^{nw+1} - p_k - (d - e))^+, \dots, d - e\}$.

The recursive relation skips index k (since job J_k is not a *WI-job*. At stage l ($l \neq k$), job J_l is either placed in machine M^1 or M^2 .

Outline for Algorithm 2:

The algorithm for the two parallel machine case is given in Algorithm 2 in Appendix B. Conceptually, we proceed as follows:

1. We first calculate cumulative weights for *E-jobs* and *T-jobs* according to the α and β values. The cumulative weights $(0, 0, \alpha, 2\alpha, 2\alpha, \dots, \beta, \beta, 2\beta, 2\beta, \dots)$ are generated

in order of increasing value, but are indexed from n in decreasing order: $cw_n = 0$, $cw_{n-1} = 0$, $cw_{n-2} = \min(\alpha, \beta)$, etc. until cw_{nw} . These cumulative weights are assigned to the corresponding *E-jobs* and *T-jobs*. For the *W-jobs*, we either need to use cw_{nw+2} and cw_{nw+1} or cw_{nw+1} and cw_{nw} , as will be shown below. Recall that according to Lemma 9, there exists an optimal schedule where the number of *WI-jobs* is either $nw - 1$ or nw .

If the number of *WI-jobs* is $nw - 1$, then we have by Lemma 10 that $n(W) = nw + 1$ and furthermore, $E \cup T = \{J_{nw+2}, \dots, J_n\}$ holds. The contribution of the *E-jobs* and *T-jobs* to the objective function is

$$\sum_{j=nw+2}^n cw_j * p_j.$$

If the number of *WI-jobs* is nw , we have $n(W) \leq nw + 2$ and furthermore that $E \cup T = \{J_{nw+3}, \dots, J_n\}$ holds. Their contribution to the objective function is

$$\sum_{j=nw+3}^n cw_j * p_j.$$

We examine each schedule with nw *WI-jobs* in Step 2 and each schedule with $nw - 1$ *WI-jobs* in Step 3. Among these schedules, there exists one which has the optimal objective value.

2. We use the Dynamic Program to examine all schedules which have nw *WI-jobs*. According to Lemma 10, the *WI-jobs* satisfy $WI \subset \{J_1, \dots, J_{nw+1}\}$. Due to the number of *WI-jobs*, we need to take out one job. Accordingly, for each job J_k where $k \in \{1, \dots, nw + 1\}$, we need to consider $WI = \{J_1, \dots, J_{nw+1}\} - \{J_k\}$. Among the non-*WI-jobs*, jobs J_k and J_{nw+2} are the next two smallest jobs and will be scheduled next to the *WI-jobs*. Due to symmetry, we only need to examine those situations where

$$\sum_{J_j \in WI^1 \cap \{J_1, \dots, J_{nw}\}} p_j \leq \sum_{J_j \in WI^2 \cap \{J_1, \dots, J_{nw}\}} p_j$$

We examine both cases for jobs J_k and J_{nw+2} :

- (a) Job J_k is scheduled on machine M^1 and Job J_{nw+2} is scheduled on machine M^2 :

$$J_k \in JS^1 \text{ and } J_{nw+2} \in JS^2$$

- (b) Job J_{nw+2} is scheduled on machine M^1 and Job J_k is scheduled on machine M^2 :

$$J_{nw+2} \in JS^1 \text{ and } J_k \in JS^2$$

Please note that we do not need to consider the cases

- (a) $J_k \in JS^1$ and $J_{nw+2} \in JS^1$ or

- (b) $J_{nw+2} \in JS^2$ and $J_k \in JS^2$.

Subsequently, we calculate the associated cost and keep

- (a) the minimum objective value
- (b) the cumulative processing of the *WI-jobs* in machine M^1
- (c) the machine to which job J_k is allocated

in order to be able to reconstruct the schedule.

3. Next, we examine all schedules with $nw - 1$ *WI-jobs* similarly to the situation in Step 2 and according to Lemma 10.
4. If there is an optimal solution with nw *WI-jobs*, we calculate the location of the common due window such that the starting times for all jobs are nonnegative and either $S_n = 0$ or $S_{n-1} = 0$ holds.
5. This step is similar to Step 4 and concerns the case that the optimal solution has $nw - 1$ *WI-jobs*.

Theorem 3 Algorithm 2 produces a schedule which minimizes $\sum_{j=1}^n (\alpha * E_j + \beta * T_j)$ in time $O(n^2 * (d - \epsilon))$.

Proof. The steps are according to the outline. According to Lemma 9, we know that there is an optimal schedule σ which satisfies $nw - 2 < n(WI(\sigma^*)) \leq nw$. According to

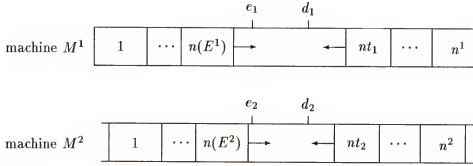


Figure 3.4. Positions of the *E-jobs* and *T-jobs* where $nt_i = n_i - n(T^i) + 1$

Lemma 10, the algorithm considers all the sets WI which are necessary to find an optimal solution. For each of those sets, the algorithm examines all possible workloads for machine M^1 due to *WI-jobs* (and thereby also all possible workloads for machine M^2 due to *WI-jobs*). According to Lemma 7, the contribution of jobs $J_1, \dots, J_{n(WI)+2}$ toward the objective function value is calculated. And according to Lemma 5, the total cost is obtained by adding the cost for jobs $\{J_{n(WI)+3}, \dots, J_n\}$ according to Step 1.

The location of the due window is determined in such a way that all starting times are nonnegative and one of the jobs is processed starting at 0.

The positioning of the *E-jobs* and *T-jobs* is illustrated in Figure 3.4.

Concerning the time complexity, consider the following: Step 1 is of linear complexity in time, since $n - (nw - 1)$ cumulative weights are determined and the contribution of jobs $J_{nw+2}, \dots, J_n$ or $J_{nw+3}, \dots, J_n$ requires a linear time effort. The calculation of A^{nw} and A^{nw-1} also involves linear effort.

The results of the dynamic program are utilized in Steps 2 and 3 and contribute an effort of order $O(n^2 * (d - e))$. This is due to the fact that $nw + 1$ different dynamic programs are used ($k = 1, \dots, nw + 1$) and that the state space is $(nw + 1) * (d - e)$. In Step 2, we have a *FOR* loop for each $k \in \{1, \dots, nw + 1\}$. Note that in the *FOR* loop for t , the range for time t is $Upper - Lower + 1 \leq d - e$. We have a constant number of operations for each value of t . The total effort for Step 2 and Step 3 is therefore bounded by $O(n^2 * (d - e))$.

The Steps 4 and 5 are of linear complexity, since the summation of PE^i and PT^i for $i \in \{1, 2\}$ involves at most $O(4 * n)$ operations, the determination of the starting times has also linear effort and the remaining parts have constant effort.

Therefore, the overall complexity is determined by Algorithm 1 and is $O(n^2 * (d - \epsilon))$.

QED

Heuristic for Two Machine Case

In the following, we present a heuristic with linear time complexity and provide an expression for the absolute error bound. The heuristic is based on a greedy strategy to place as many jobs within each window as possible. Recall that the pseudopolynomial time complexity in Algorithm 2 stems from the determination of the maximal number of W -jobs using Algorithm 1.

Algorithm 3: Heuristic

Step 1. Same as in Algorithm 2

Step 2.-5. Apply a greedy strategy:

$$n(WI^1) = \max\{l : \sum_{i=1}^l p_j \leq (d - \epsilon)\}$$

$$n(WI^2) = \max\{l : \sum_{i=n(WI^1)+1}^{n(WI^1)+l} p_j \leq (d - \epsilon)\}$$

$$n(WI) = n(WI^1) + n(WI^2)$$

Do all the calculations as in Steps 2.-5. for the case $n(WI)$ WI -jobs and

$$t = \sum_{j=1}^{n(WI^1)} p_j.$$

Note that the heuristic generates a schedule which satisfies *Properties 1-3*. In the following, we first provide an upper bound for the difference of two consecutive cumulative weights.

Lemma 11 Let $\{cw_{n-2*i} : 0 \leq i \leq \lfloor (n - nw)/2 \rfloor\}$ be the $\lfloor (n - nw)/2 + 1 \rfloor$ lowest numbers of the set $\{k * \alpha : k \geq 0 \text{ and } k \text{ is integer}\} \cup \{k * \beta : k \geq 0 \text{ and } k \text{ is integer}\}$ such that $cw_m \leq cw_{m-2}$ holds for $m = n - 2 * (\lfloor (n - nw)/2 \rfloor + 1), \dots, n$. In addition, let $cw_{n-2*i-1} =$

cw_{n-2+i} for $i = 0, \dots, \lfloor (n - nw)/2 \rfloor - 1$. In other words, these weights $cw_n, \dots, cw_{nw}$ are identical to the ones which are generated in Step 1 of Algorithm 2.

Then we have

$$0 \leq cw_m - cw_{m+1} \leq \min\{\alpha, \beta\} \quad \text{for } nw \leq m \leq n - 1.$$

Proof. By definition, we have $0 \leq cw_m - cw_{m+1}$ for $nw \leq m \leq n - 1$. Now assume that $cw_m - cw_{m+1} > \min\{\alpha, \beta\}$ holds. Let $\gamma = \min\{\alpha, \beta\}$ and define $cw = \min\{k * \gamma : k * \gamma > cw_{m+1}\}$. By definition of cw , we have $cw - \gamma \leq cw_{m+1}$ and by $cw_{m+1} + \gamma < cw_m$, we have $cw < cw_m$. Since cw belongs to the set $\{k * \alpha : k \geq 0 \text{ and } k \text{ is integer}\} \cup \{k * \beta : k \geq 0 \text{ and } k \text{ is integer}\}$, this provides a contradiction to the definition of the numbers $cw_n, \dots, cw_{nw}$. **QED**

Lemma 12 *The Complexity of Algorithm 3 is linear ($O(n)$), provided that the jobs are indexed in SPT order. Otherwise, the complexity is $O(n * \log(n))$.*

The absolute error is at most $\min(p_{nw+2}, d - \epsilon) * \min(\alpha, \beta)$ for $nw + 2 \leq n$ and 0 otherwise.

Proof. Concerning the complexity, note that the *W-jobs* are heuristically determined in linear time. Since only one case for $n(WI)$ and only one value for the workload of *WI-jobs* on machine M^1 is considered, the heuristic requires an effort which is linear in time. Further, sorting of the indices results in a total effort of $O(n * \log(n))$.

Concerning the absolute error, note the following: According to the definition of nw , the heuristic will find at least $nw - 1$ *WI-jobs*. The heuristic will therefore obtain 0 cost when the number of jobs satisfies $n \leq nw + 1$. In the following, we therefore only need to consider the case $n \geq nw + 2$. Also, the heuristic will produce a schedule which satisfies *Properties 1-3* and Equation (3.6). Furthermore, the *E-jobs* and *T-jobs* will be optimally scheduled according to the cumulative weights.

According to Lemma 9, every optimal schedule σ^* which satisfies Equation (3.6) also satisfies Equation (3.11), i.e., $nw - 1 \leq n(WI(\sigma^*)) \leq nw$ for this two machine case. In

order to compare the cost of the schedule obtained by the heuristic to the optimal objective value, we first consider two schedules σ_0 and σ_1 which satisfy *Properties 1-3*, Equation (3.6), and further $n(WI(\sigma_0)) = nw$ and $n(WI(\sigma_1)) = nw - 1$. Furthermore, we require that for both schedules, the *E-jobs* and *T-jobs* are optimally scheduled according to their cumulative weights, i.e., if we have $J_j, J_l \in E(\sigma_i) \cup T(\sigma_i)$ and $p_j < p_l$, then we must have $cw(J_j, \sigma_i) \geq cw(J_l, \sigma_i)$ for $i = 1, 2$.

According to the different cases in Lemma 10, we can write the objective function value of schedule σ_0 in each case where $n(WI) = nw$ holds as follows

$$\begin{aligned} cost(\sigma_0) &= cw_{nw+1} * (p_{k0} + T_1 - (d - \epsilon)) + cw_{nw+2} * (p_{nw+2} + T_2 - (d - \epsilon)) \\ &\quad + \sum_{j=nw+3}^n cw_j * p_j \end{aligned}$$

where $k0 \in \{1, \dots, nw + 1\}$, $T_1 + T_2 + p_{k0} = A^{nw+1}$, and $T_1 \leq T_2 \leq d - \epsilon$ hold. Due to optimality and $cw_{nw+1} \geq cw_{nw+2}$, we can restrict ourselves to the case that $T_1 \leq T_2$ holds. The expression above represents the case that $WI(\sigma_0) = \{J_1, \dots, J_{nw+1}\} - \{J_{k0}\}$, $\sum_{J_j \in WI^1(\sigma_0)} p_j = T_1$, and $\sum_{J_j \in WI^2(\sigma_0)} p_j = T_2 (\geq T_1)$ hold where $\{i1, i2\} = \{1, 2\}$. Due to optimality, $cw_{nw+1} \geq cw_{nw+2}$, and $p_{k0} \leq p_{nw+2}$, we only need to consider the case that job J_{k0} is scheduled on the same machine which has a load of T_1 processing time units due to *WI-jobs*.

Likewise, according to the different cases in Lemma 10, we can write the objective function value of schedule σ_1 in each case where $n(WI) = nw - 1$ holds as follows

$$\begin{aligned} cost(\sigma_1) &= cw_{nw} * (p_{k1} + T_3 - (d - \epsilon)) + cw_{nw+1} * (p_{nw+1} + T_4 - (d - \epsilon)) \\ &\quad + \sum_{j=nw+2}^n cw_j * p_j \end{aligned}$$

where $k1 \in \{1, \dots, nw\}$, $T_3 + T_4 + p_{k1} = A^{nw}$, and $T_3 \leq T_4 \leq d - \epsilon$ hold.

According to the generation of the cumulative weights (which is independent of schedule σ_0 or σ_1), we have either $cw_{nw} = cw_{nw+1}$ or $cw_{nw+1} = cw_{nw+2}$:

1. $cw_{nw} = cw_{nw+1}$:

In this case, we have

$$\begin{aligned}
& cost(\sigma_0) - cost(\sigma_1) \\
&= \{cw_{nw+1} * (p_{k0} + T_1 - (d - e)) + cw_{nw+2} * (p_{nw+2} + T_2 - (d - e))\} \\
&- \{cw_{nw+1} * (A^{nw+1} - 2 * (d - e)) + cw_{nw+2} * p_{nw+2}\} \\
&= cw_{nw+1} * ((d - e) - (A^{nw+1} - p_{k0} - T_1)) + cw_{nw+2} * (T_2 - (d - e)) \\
&= (cw_{nw+1} - cw_{nw+2}) * ((d - e) - T_2)
\end{aligned}$$

Please note that $0 \leq ((d - e) - T_2) \leq \min((d - e), p_{nw+2})$ holds due to $d - e \geq T_2$, $(d - e) - T_2 \leq (d - e)$ and $T_2 + p_{nw+2} \geq (d - e)$ (i.e., $(d - e) - T_2 \leq p_{nw+2}$). Using Lemma 11, we conclude

$$0 \leq cost(\sigma_0) - cost(\sigma_1) \leq (\min(\alpha, \beta)) * \min((d - e), p_{nw+2}).$$

2. $cw_{nw+1} = cw_{nw+2}$:

In this case, we have

$$cost(\sigma_0) - cost(\sigma_1) = (cw_{nw+1} - cw_{nw}) * (p_{k1} + T_3 - (d - e))$$

Note that $cost(\sigma_0) - cost(\sigma_1) \leq 0$, since we have $(cw_{nw+1} - cw_{nw}) \leq 0$ and $(p_{k1} + T_3 - (d - e)) \geq 0$. Similarly, we have $0 \leq p_{k1} + T_3 - (d - e) \leq p_{k1} \leq p_{nw+2}$ due to $p_{k1} + T_3 \geq d - e$ and $T_3 \leq d - e$. Since schedule σ_0 has nw WI-jobs, we must have $p_{nw} \leq d - e$. Since $k1 \leq nw$ holds, we conclude $p_{k1} \leq d - e$. Using Lemma 11, we finally have

$$-(\min(\alpha, \beta)) * \min((d - e), p_{nw+2}) \leq cost(\sigma_0) - cost(\sigma_1) \leq 0$$

We can write for both cases ($cw_{nw} = cw_{nw+1}$ and $cw_{nw+1} = cw_{nw+2}$)

$$|cost(\sigma_0) - cost(\sigma_1)| \leq (\min(\alpha, \beta)) * \min((d - e), p_{nw+2}).$$

Now, if the heuristic and the optimal schedule have a different number of *WI-jobs*, we have by the above considerations that the absolute error bound is $(\min(\alpha, \beta)) * \min((d - \epsilon), p_{nw+2})$. If the heuristic and the optimal schedule have the same number of *WI-jobs*, the same cumulative weight will be applied to jobs $J_{n(WI)+2}, \dots, J_n$. Only due to the situation where there are 2 *W-jobs* which are not *WI-jobs*, there could be a difference due to the smaller of these two jobs. Again, an upper bound for this difference is $(\min(\alpha, \beta)) * \min((d - \epsilon), p_{nw+2})$.

Therefore, in all cases, the absolute error is at most $(\min(\alpha, \beta)) * \min((d - \epsilon), p_{nw+2})$.

QED

3.5 Generalization for the m machine case

According to Lemma 5, the cumulative weight of *E-jobs* and *T-jobs* is easily determined by generating the cumulative weights in ascending order and assigning the jobs in LPT order to them. Thus, to develop a polynomial heuristic for the general machine case, we only need to determine the *W-jobs*. The completion of the schedule with *E-jobs* and *T-jobs* is straightforward. A generic description of the algorithm is as follows:

Algorithm 4: *Heuristic for the parallel machine case*

Step 1. Load machines $M^1, M^2, \dots, M^m$ with *W-jobs* according to a greedy fashion.

$index = 1$ (index of next job to be considered)

FOR $i = 1$ TO m DO

$\{W^i(\sigma^H) \leftarrow \emptyset$

$P_i = 0$ (time consumed by these *W-jobs*)

WHILE $(P_i + p_{index} < d - \epsilon)$ DO

$\{P_i \leftarrow P_i + p_{index}$

$W^i(\sigma^H) \leftarrow W^i(\sigma^H) \cup \{J_{index}\}$

$index \leftarrow index + 1\}$

$$W^i(\sigma^H) \leftarrow W^i(\sigma^H) \cup \{J_{index}\}$$

$$index \leftarrow index + 1\}$$

Step 2. Generate cumulative weights $cw_n^m, \dots, cw_{nw}^m$ in nondecreasing order similarly to Step 1 in Algorithm 2 while each weight is repeated m times. Assign the remaining jobs (the non- W -jobs) accordingly.

Step 3. After determining the W -jobs, the E -jobs, and the T -jobs for each machine, the location of the common due window and the starting times for the jobs can be determined such that the starting times are nonnegative, *Properties 1-3* are satisfied, and that there is one machine which starts processing at time 0.

Remark 8 In Step 2, there are groups of m identical cumulative weights since there are m machines. Multiple solutions exist and the jobs can be assigned to any machine as long as their position results in the same cumulative weight. Therefore, a secondary criterion can be used to break these ties. See also Emmons [38].

Theorem 4 The absolute error for the heuristic schedule σ^H above is

$$m * \min(p_{max}, d - \epsilon) * \min(\alpha, \beta)$$

The relative error bound vanishes as the number of jobs increases. That is,

$$\lim_{n \rightarrow \infty} \frac{cost(\sigma^H) - cost(\sigma^*)}{cost(\sigma^*)} = 0$$

where $cost(\sigma^H)$ is the objective value for schedule σ^H and $cost(\sigma^*)$ is the optimal objective value. It is assumed that m is fixed and that $cost(\sigma^*)$ for the relative error to be defined.

Proof. Note that the heuristic schedule σ^H satisfies Equation (3.6) according to Step 1 in Algorithm 4 and *Properties 1-3* according to Step 3. Therefore, we have $nw - m < n(WI(\sigma^H)) \leq nw$ according to Lemma 9.

Let σ^* be an optimal schedule which satisfies *Properties 1-3* and Equation (3.6). Therefore, by Lemma 9, we conclude that $|n(WI^H) - n(WI(\sigma^*))| \leq m - 1$ holds.

For each job J_j , let p_j^H and p_j^* be the units of processing time of job J_j which are processed during the due window in schedules σ^H and σ^* , respectively. Note that we have

$$\sum_{j=1}^n p_j^H = m * (d - \epsilon) = \sum_{j=1}^n p_j^* \quad (3.12)$$

by definition of p_j^H and p_j^* , and by $\text{cost}(\sigma^*) > 0$. Using Lemma 5 and the definition of p_j^H and p_j^* , the difference in the objective value can be written as follows:

$$\begin{aligned} \text{cost}(\sigma^H) - \text{cost}(\sigma^*) &= \sum_{i=1}^m \{ \{ \sum_{J_j \in JS^i(\sigma^H)} cw(J_j, \sigma^H) * p_j \} - cw(W^i, \sigma^H) * (d - \epsilon) \} \\ &\quad - \sum_{i=1}^m \{ \{ \sum_{J_j \in JS^i(\sigma^*)} cw(J_j, \sigma^*) * p_j \} - cw(W^i, \sigma^*) * (d - \epsilon) \} \\ &= \sum_{i=1}^m \{ \{ \sum_{J_j \in JS^i(\sigma^H)} cw(J_j, \sigma^H) * (p_j - p_j^H) \} \} \\ &\quad - \sum_{i=1}^m \{ \{ \sum_{J_j \in JS^i(\sigma^*)} cw(J_j, \sigma^*) * (p_j - p_j^*) \} \} \\ &= \sum_{j=1}^n \{ cw(J_j, \sigma^H) * (p_j - p_j^H) - cw(J_j, \sigma^*) * (p_j - p_j^*) \} \end{aligned}$$

Now note that jobs with $p_j^H = p_j^* = 0$ do not make a nonzero contribution to the cost difference since a job J_j with $p_j^H = p_j^* = 0$ is either an *E-job* or a *T-job* and has the same cumulative weight in both schedules. Define the index-set IS by

$$IS = \{J_j : p_j^H > 0 \text{ or } p_j^* > 0\}.$$

Accordingly, we can write

$$\text{cost}(\sigma^H) - \text{cost}(\sigma^*) = \sum_{J_j \in IS} \{ cw(J_j, \sigma^H) * (p_j - p_j^H) - cw(J_j, \sigma^*) * (p_j - p_j^*) \}.$$

Further note that for all jobs J_j with $p_j > d - \epsilon$, we have $cw(J_j, \sigma^H) = cw(J_j, \sigma^*)$ since their cumulative weight solely depends on the number of *E-jobs* preceding or the number of *T-jobs* succeeding them. Therefore, we can write

$$\begin{aligned} \text{cost}(\sigma^H) - \text{cost}(\sigma^*) &= \sum_{J_j \in IS} \{ \quad cw(J_j, \sigma^H) * (\min(d - \epsilon, p_j) - p_j^H) \\ &\quad - \quad cw(J_j, \sigma^*) * (\min(d - \epsilon, p_j) - p_j^*) \}. \end{aligned} \quad (3.13)$$

Now consider the cumulative weights in the heuristic which are applied to the first m non- WI -jobs. We have by Lemma 6 and by Lemma 11 that these differ by at most $\min(\alpha, \beta)$ for any two machines. The same statement also holds for the first m non- WI -jobs in the optimal schedule. Due to the fact that the number of WI -jobs in schedules σ^* and σ^H differs by at most $m - 1$, the cumulative weights of the first m non- WI -jobs of σ^* and of σ^H differ by at most $\min(\alpha, \beta)$.

After rearranging the terms in Equation (3.13), we obtain

$$\begin{aligned}
 & cost(\sigma^H) - cost(\sigma^*) \\
 & \leq \sum_{J_j \in IS} \min(\alpha, \beta) * (\min(d - \epsilon, p_j) - p_j^H) \\
 & = \sum_{J_j \in IS} \min(\alpha, \beta) * (\min(d - \epsilon, p_j) - p_j^*) \\
 & = \min(\alpha, \beta) * \{ (\sum_{J_j \in IS} \min(d - \epsilon, p_j)) - m * (d - \epsilon) \}
 \end{aligned}$$

According to the definition of IS , we have $IS = W(\sigma^*) \cup W(\sigma^H)$. The quantity $(\sum_{J_j \in IS} \min(d - \epsilon, p_j)) - m * (d - \epsilon)$ is the total processing time by which the W -jobs of either schedule σ^* or σ^H exceed all the due windows. This is bounded above by the processing time units of the m largest W -jobs and we have

$$(\sum_{J_j \in IS} \min(d - \epsilon, p_j)) - m * (d - \epsilon) \leq m * \min((d - \epsilon), p_{max}).$$

Therefore, the absolute error is bounded above by

$$m * \min(p_{max}, d - \epsilon) * \min(\alpha, \beta).$$

Now we consider the relative error. We do this only for schedules with positive objective value. For this, we produce a rough lower bound for any instance with positive cost in an optimal schedule. Clearly, for all non W -jobs except the largest m ones, the cumulative costs are $1 * \alpha, 2 * \alpha, \dots, 1 * \beta, 2 * \beta, \dots$. Therefore, all E -jobs and T -jobs, except the m largest ones, have a cumulative weight which is at least $\min(\alpha, \beta)$. Therefore, we have for the optimal objective value:

$$cost(\sigma^*) \geq \min(\alpha, \beta) \sum_{j=nw+m+1}^{n-m} p_j$$

Using the expression for the absolute error, we have

$$\begin{aligned}
\frac{cost(\sigma^H) - cost(\sigma^*)}{cost(\sigma^*)} &\leq \frac{m * \min(p_{max}, d-e) * \min(\alpha, \beta)}{cost(\sigma^*)} \\
&\leq \frac{m * \min(p_{max}, d-e) * \min(\alpha, \beta)}{\min(\alpha, \beta) \sum_{j=n(WI^*)+m+1}^{n-m} p_j} \\
&\leq \frac{m * (d-e)}{((n-m) - (n(WI^*) + m + 1) - 1)} \\
&< \frac{m * (d-e)}{n - (2 * (m+1) + n(WI^*))} \\
&\leq \frac{m * (d-e)}{n - (2 * (m+1) + m * (d-e))}
\end{aligned}$$

The last inequality is derived from the assumption that jobs have integral processing times, i.e., $p_j \geq 1$ and from the definition of nw . Therefore, if the number of machines is a given parameter, the relative error approaches zero as the number of jobs increases:

$$\lim_{n \rightarrow \infty} \frac{cost(\sigma^H) - cost(\sigma^*)}{cost(\sigma^*)} = 0$$

QED

3.6 Summary

In this chapter, we have extended the concept of due window scheduling to parallel machines. It has been shown that even the decision variable unit weights for two machines is NP-hard. This justifies the use of dynamic programming and heuristics. A dynamic programming algorithm was provided for the two machine case. The non-polynomial portion of the algorithm is the determination of *WI-jobs*. A simple heuristic for this part has been used to achieve polynomial complexity. The heuristic version for parallel machines is especially suited for multi-criteria optimization due to multiple common cumulative weights. If the number of jobs is reasonably large, the relative error in the objective function tends to be small since the relative error vanishes as the number of jobs increases.

CHAPTER 4

OPTIMAL CONTROL OF A PRODUCTION SYSTEM WITH PERIODIC MAINTENANCE

4.1 Introduction

Production scheduling is an important area both for researchers and for practitioners, which is reflected in the vast amount of existing literature. For a survey, see, for example, Lawler et al. [75]. A large section of the existing literature approaches the problem with the assumption that the machines are reliable during the period of consideration. Machine breakdown and machine maintenance oftentimes are not considered, mostly for simplification.

Since the early 1980s, there has been an increased interest in developing production flow control models for systems with unreliable machines. Kimemia and Gershwin [66] have studied a flexible manufacturing system of unreliable machines where they used a production flow model. In the model of Kimemia and Gershwin, the machine state is governed by a finite state Markov chain, and part movement is assumed to be continuous. Their main contribution is the qualitative characterization of an optimal control structure. They introduced the notion of *hedging point*. The hedging point is a production surplus level which is utilized to hedge against random machine failures. The control policy, also referred to as hedging policy or threshold policy, operates as follows: There is a target inventory level (referred to as threshold level or hedging level), and the system tries to reach this level as soon as possible. When the actual inventory is below the target level, production is at full capacity. When the actual inventory is above the target, the machine is shut down. When the target inventory level is reached, the production rate equals the demand. Further, based on assumptions concerning the frequency of events (demand changes occur less frequently than machine failures and operation times are small compared to the time

between failures), they have proposed a multi-level hierarchical control algorithm. The top level is used for long-term decisions such as the selection of the hedging point, the next level is used to control the production flow and the routing. At the lowest level of control are part-dispatch policies.

Subsequently, the concept of modeling production as continuous flow, the concept of hedging point, and the concept of a hierarchical control algorithm have been refined and applied to similar problem settings. Gershwin, Akella, and Choong [49] have improved the algorithms for the various levels of the hierarchical control. Akella, Choong, and Gershwin [4] have studied the performance of this model for an IBM printed circuit card assembly line. They have concluded that a hierarchically structured policy such as the one given in [49] is very effective in scheduling a FMS and that it can achieve high output with low WIP and can cope with changes and disturbances. Recently, Bai and Gershwin have studied finite buffer capacity models in three settings. They have considered one-part-type systems in [14], multiple-part-type systems in [15], and reentrant systems in [16]. Sharifnia, Caramanis, and Gershwin [101] and Srinivatsan and Gershwin [103] have studied unreliable systems with non-negligible setup times.

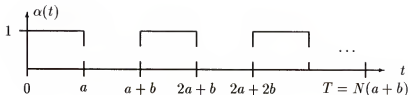
It is difficult to obtain a closed-form analytical solution to the optimal production flow control problem for a system which consists of unreliable machines. In most of the previous research in this area, only qualitative characteristics and/or heuristic solutions were obtained. In few cases, researchers have succeeded in deriving the solution to optimality. There are two papers found in the literature which deal with similar problems as the one in this study and present closed-form optimal solutions. The first one is by Akella and Kumar [5] where the authors considered a manufacturing system consisting of one unreliable machine which produces one part-type. They analytically derive an exact solution and show that the optimal control policy is a hedging policy.

The second reference where a control problem has been solved to optimality is by Hu and Xiang [60]. In that paper, the authors have considered a system consisting of one machine

where the operative time is deterministic and the length of the downtime is random. Their work focuses on the steady state for which a switching curve policy is presented. It is characterized by a single parameter which corresponds to a target inventory level. It is to be reached at the end of each uptime interval. The difference to the hedging policy is that the system maintains zero inventory level as long as possible. Before the end of each uptime interval, production proceeds at the maximum rate in order to reach the target inventory level at the end of the uptime interval. Since the length of the uptime interval is deterministic, the inventory does not need to be built up in advance, whereas the hedging policy requires that the system maintains the target inventory level at all times, since in that case, the length of the uptime interval is random.

In this chapter, we consider a single-machine system where maintenance is performed periodically. When the machine is undergoing maintenance, the system is shut down and cannot produce. Both the uptime interval and the maintenance interval are deterministic. In contrast, Akella and Kumar [5] model both the uptime and the downtime interval as intervals of random length and Hu and Xiang [60] model the downtime interval with random length. In addition, the solutions in [5] and [60] are based on the Bellman equation, while our solution is obtained via Pontryagin's Minimum Principle. We will see that the optimal policy for our problem is of the type of switching curve policy rather than hedging policy due to the deterministic length of machine uptime. In addition to the steady state solutions, we will also discuss the transient solution.

We define the general N -cycle problem in Section 4.2. In Section 4.3, we solve the one-cycle problem using Pontryagin's Minimum Principle. The results for the one-cycle problem are used to construct the optimal solution to the two-cycle problem in Section 4.4. The original N -cycle problem is solved in Section 4.5. The steady state and the transient state is discussed in Section 4.6. Concluding remarks are given in Section 4.7.

Figure 4.1. Machine Status $\alpha(t)$.

4.2 Problem Definition

We consider a system consisting of one machine which produces one part-type. The demand rate is constant and is denoted by d . The machine requires periodical maintenance. During the maintenance, the machine is shut down and cannot produce. The time between maintenances is constant, denoted by a , and the maintenance time is also constant, denoted by b .

We denote the *machine status* by a binary variable $\alpha(t)$ which equals 0 if maintenance is performed at time t and 1 if the machine is operative at time t . Without loss of generality, we assume $\alpha(0) = 1$.

We consider a finite planning horizon from 0 to $T = N(a+b)$, where N is a given integer and is greater or equal to 1. For N cycles, the machine status is governed by

$$\alpha(t) = \begin{cases} 1, & \text{if } t \in [(j-1)(a+b), (j-1)(a+b) + a], \quad j = 1, \dots, N, \\ 0, & \text{if } t \in ((j-1)(a+b) + a, j(a+b)), \quad j = 1, \dots, N. \end{cases} \quad (4.1)$$

Figure 4.1 illustrates the change of machine status. We shall notice that the time interval of operation $[0, a]$ is closed and the maintenance interval $(a, a+b)$ is open.

The *control variable* is the production rate $u(t)$. A production control $u(t)$ is *admissible* if

$$u(t) \in [0, U\alpha(t)], \quad \text{for } t \in [0, T], \quad (4.2)$$

where U is the machine capacity or maximum production rate.

For the sake of definiteness, we restrict ourselves to right-continuous production controls, i.e., production controls which satisfy

$$u^*(\tau) = \lim_{t \rightarrow \tau, t > \tau} u^*(t)$$

for points of discontinuity.

We require that the demand be feasible. The system is capable of meeting the demand if $aU \geq (a+b)d$, or equivalently,

$$d \leq \frac{a}{a+b}U. \quad (4.3)$$

We do not consider the case $d > \frac{a}{a+b}U$. Such a case is of little interest from a control perspective, since the demand is greater than the system capacity.

The *production surplus* $x(t)$ is defined as the cumulative difference between total production and total demand. It satisfies

$$\frac{dx(t)}{dt} = u(t) - d.$$

For a given initial state $x(0) = x_0$, we have

$$x(t) = x_0 + \int_0^t (u(t) - d)dt, \quad \text{for } t \in [0, T]. \quad (4.4)$$

Our objective is to keep the cumulative production close to the cumulative demand. In terms of the production surplus $x(t)$, our objective is to keep the production surplus $x(t)$ close to zero. The following differentiable and strictly convex penalty function $g(x(t))$ reflects this objective,

$$g(x(t)) = x^2(t). \quad (4.5)$$

We integrate the penalty function over the time interval $[0, T]$ to obtain the performance measure. Both the trajectory and the performance measure depend on the production control. We have

$$J(u) = \int_0^T x^2(t)dt. \quad (4.6)$$

We now can formally define the optimal production control problem:

Problem P_N : Find an admissible production control $u_N^*(t)$ that causes the system

$$\frac{dx(t)}{dt} = u(t) - d$$

to follow a trajectory $x^*(t)$ which minimizes the performance measure

$$J_N(u) = \int_0^T x^2(t) dt \quad (4.7)$$

where $T = N(a + b)$ is a given constant. The initial state $x(0)$ is fixed and the final state $x(T)$ is free.

Here and elsewhere in this chapter, we use a subscript to indicate the number of cycles involved in a problem. For example, $J_1^*(.)$ and $u_1^*(.)$ are the optimal cost and control policy for Problem P_1 ; $J_2^*(.)$ and $u_2^*(.)$ are the optimal cost and control policy for Problem P_2 ; etc. However, the subscript is omitted in the derivation when there is no possible ambiguity. It is used only when comparison is made for results of various number of cycles.

In the following sections, we solve for the optimal solution to the problem above. In Section 4.3, we consider the one-cycle problem which is obtained by letting $N = 1$ in the definition of Problem P_N (see Equation (4.7)). We then use the results from the one-cycle problem to solve the two-cycle problem in Section 4.4 and finally extend the results to solve the N -cycle problem for $N \geq 3$ in Section 4.5.

4.3 A One-Cycle Problem and its Optimal Solution

In the previous section, we formulated an optimal control problem for the system under consideration. In this section, we study a special case of such an optimal control problem which has only one up-down cycle ($N = 1$). This one-cycle problem will be solved to optimality via Pontryagin's Minimum Principle. The results will be extended to the two-cycle problem in the next section.

For $t \in (a, a + b)$, since the machine is undergoing maintenance, there is no control decision to make, i.e., we have

$$\begin{aligned} u(t) &= 0, & \text{for } t \in (a, a + b), \\ x(t) &= x(a) - d(t - a), & \text{for } t \in (a, a + b). \end{aligned} \quad (4.8)$$

The cost during the interval $(a, a + b)$ can therefore be expressed in dependence of $x(a)$. We introduce the terminal cost function $h_1(x(a))$ by defining

$$h_1(x(a)) = \int_a^{a+b} (x(a) - d(t - a))^2 dt = \int_0^b (x(a) - dt)^2 dt. \quad (4.9)$$

Notice that $h_1(x(a))$ satisfies

$$\frac{d}{dx(a)} h_1(x(a)) = \int_0^b 2(x(a) - dt) dt = 2x(a)b - db^2 = 2(x(a) - d\frac{b}{2})b = 2x(a + \frac{b}{2})b. \quad (4.10)$$

The control problem for the one-cycle case can be restated as follows.

Problem P_1 : Determine an admissible control $u_1^*(t)$, i.e., $u_1^*(t) \in [0, U]$, for $t \in [0, a]$, that causes the system

$$\frac{dx(t)}{dt} = u(t) - d$$

to follow a trajectory $x^*(t)$ which minimizes the performance measure

$$J_1(u) = \int_0^a x^2(t) dt + h_1(x(a)). \quad (4.11)$$

The boundary conditions are $x(0) = x_0$ and $x(a)$ free.

For Problem P_1 , the Hamiltonian (see Pontryagin [86], p. 18 and Kirk [67], p. 188) is given by

$$\mathcal{H}(x(t), u(t), p(t)) = x^2(t) + p(t)[u(t) - d], \quad \text{for } t \in [0, a],$$

where the function $p(t)$ is referred to as costate function.

According to Pontryagin's Minimum Principle, necessary conditions for an optimal control law $u^*(t)$ (see Pontryagin [86], p. 69 and Kirk [67], p. 233) are

$$\begin{aligned} \frac{dx^*(t)}{dt} &= \frac{\partial \mathcal{H}}{\partial p}(x^*(t), u^*(t), p^*(t)); \\ \frac{dp^*(t)}{dt} &= -\frac{\partial \mathcal{H}}{\partial x}(x^*(t), u^*(t), p^*(t)); \end{aligned}$$

$$\mathcal{H}(x^*(t), u^*(t), p^*(t)) \leq \mathcal{H}(x^*(t), u(t), p^*(t)), \quad \text{for all admissible } u(t);$$

and the boundary condition (see Kirk [67], p. 200)

$$\frac{\partial h}{\partial x}(x^*(T)) - p^*(T) = 0.$$

These evaluate in case of Problem P_1 to

$$\frac{dx^*(t)}{dt} = u^*(t) - d; \quad (4.12)$$

$$\frac{dp^*(t)}{dt} = -2x^*(t); \quad (4.13)$$

$$\begin{aligned} x^{*2}(t) + p^*(t)[u^*(t) - d] &= \min_{u(t) \in [0, U]} \{x^{*2}(t) + p^*(t)[u(t) - d]\} \\ &= x^{*2}(t) - p^*(t)d + \min_{u(t) \in [0, U]} \{p^*(t)u(t)\}; \end{aligned} \quad (4.14)$$

$$2x^*(a)b - db^2 - p^*(a) = 0. \quad (4.15)$$

Since the final time is fixed and the Hamiltonian does not explicitly depend on time, the Hamiltonian must satisfy the following condition for the optimal control $u^*(t)$, the optimal trajectory $x^*(t)$, and the costate function $p^*(t)$,

$$\mathcal{H}(x^*(t), u^*(t), p^*(t)) = c_1, \quad \text{for all } t,$$

where c_1 is a constant (see Pontryagin [86], p. 19 and Kirk [67], p. 236).

For Problem P_1 , we have

$$x^{*2}(t) + p^*(t)[u^*(t) - d] = c_1, \quad \text{for } t \in [0, a]. \quad (4.16)$$

Since there is only one term in the Hamiltonian involving $u(t)$, we can evaluate $u(t)$ according to Equation (4.14) of Pontryagin's Minimum Principle. This results in

$$u^*(t) = \begin{cases} 0, & \text{if } p^*(t) > 0, \\ U, & \text{if } p^*(t) < 0, \\ \beta, & \text{if } p^*(t) = 0, \end{cases} \quad (4.17)$$

where $\beta \in [0, U]$ is to be determined.

According to Equation (4.17), the optimal control law $u^*(t)$ is thus determined whenever we have $p^*(t) \neq 0$. In the case that roots exist for the function $p^*(t)$ within the interval $[0, a]$,

we further distinguish between *singular intervals*, *switching times*, and *switching points*. These terms are defined in the following.

An interval $[t_1, t_2]$ with $t_1 < t_2$ is called a *singular interval* if the following two conditions are satisfied:

$$\begin{aligned} \text{(i)} \quad p^*(t) &= 0 & \text{for } t \in [t_1, t_2], \quad \text{and} \\ \text{(ii)} \quad p^*(t) &= 0 & \text{for } t \in [t'_1, t'_2] \text{ with } t'_1 \leq t_1, t_2 \leq t'_2 \\ & \text{if and only if} & t'_1 = t_1 \text{ and } t'_2 = t_2. \end{aligned} \quad (4.18)$$

In case of $t_1 > 0$, the definition implies that there must exist a neighborhood to the left of t_1 in which $p^*(t)$ is nonzero, i.e., there exists an $\epsilon > 0$ such that $p^*(t) \neq 0$ for $t \in (t_1 - \epsilon, t_1)$. Similarly, in case of $t_2 < a$, the definition implies that there must exist a neighborhood to the right of t_2 in which $p^*(t)$ is nonzero, i.e., there exists an $\epsilon > 0$ such that $p^*(t) \neq 0$ for $t \in (t_2, t_2 + \epsilon)$.

A time t_0 is called a *switching time* if

$$\begin{aligned} \text{(i)} \quad p^*(t_0) &= 0, \quad \text{and} \\ \text{(ii)} \quad t_0 &\text{ is not part of any singular interval.} \end{aligned} \quad (4.19)$$

That is, t_0 is an isolated time instant at which $p^*(t_0) = 0$. We do not consider 0 and a as switching times, i.e., t_0 must satisfy $t_0 \in (0, a)$. The production surplus $x(t_0)$ is referred to as a *switching point*.

In the following, we will consider the implications for a control which has a singular interval and subsequently, we study the implications for a control which has a switching point. We will show that an optimal control will either have one singular interval and no switching point, or one switching point and no singular interval, or neither of them.

4.3.1 Singular Intervals

Now assume that there is a singular interval $[t_1, t_2] \subseteq [0, a]$, i.e.,

$$p^*(t) = 0, \quad \text{for } t \in [t_1, t_2]. \quad (4.20)$$

Using Equations (4.13) and (4.12), we conclude that the optimal trajectory and the optimal control policy must satisfy

$$x^*(t) = 0, \quad \text{for } t \in (t_1, t_2), \quad (4.21)$$

$$u^*(t) - d = 0, \quad \text{for } t \in (t_1, t_2). \quad (4.22)$$

Thus, for a singular interval, we must have $\beta = d$ in Equation (4.17). From this equation, we conclude that the optimal control $u^*(t)$ is a piecewise constant function for $t \in [0, a]$. From Equation (4.4), we have that the optimal trajectory $x^*(t)$ is a continuous and piecewise linear function for $t \in [0, a]$. This implies $x^*(t_1) = x^*(t_2) = 0$, so that Equation (4.21) can be extended from (t_1, t_2) to $[t_1, t_2]$:

$$x^*(t) = 0 \quad \text{for } t \in [t_1, t_2]. \quad (4.23)$$

Using our convention that we only consider right-continuous production rates, we can extend Equation (4.22) from (t_1, t_2) to $[t_1, t_2]$:

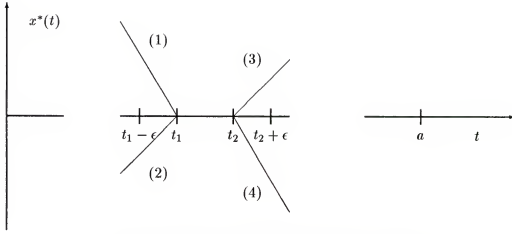
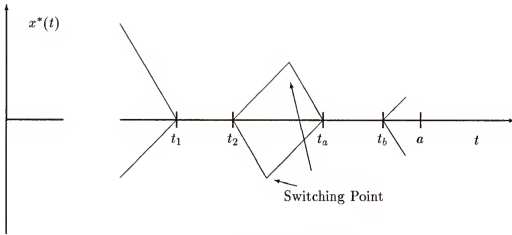
$$u^*(t) = d, \quad \text{for } t \in [t_1, t_2].$$

We now proceed to analyze the implications of having a singular interval and consider the constant c_1 in Equation (4.16). According to Equations (4.20) and (4.23), we have $c_1 = 0$, and thus,

$$x^{*2}(t) + p^{*2}(t)(u^*(t) - d) = 0, \quad \text{for } t \in [0, a].$$

Clearly, a root of the costate function $p^*(t)$ requires a zero production surplus level. In other words, if the production surplus level is nonzero at a time t ($x^*(t) \neq 0$), the time t cannot be a root of the costate function $p^*(t)$, i.e., $p^*(t) \neq 0$, and therefore, t cannot be contained in a singular interval.

Therefore, if there exists a singular interval, then during this singular interval and any other singular interval, and also at any switching point, we must have zero production

(a) Optimal trajectory in a neighborhood of $[t_1, t_2]$.

(b) Two singular intervals.

Figure 4.2. A Control with Singular Interval.

surplus level in order to satisfy Equation (4.16). That is,

$$\begin{aligned} \text{if } & \text{there is a singular interval } [t_1, t_2] \subseteq [0, a], \\ \text{then } & p^*(t) = 0 \quad \text{implies} \quad x^*(t) = 0, \text{ for } t \in [0, a]. \end{aligned} \quad (4.24)$$

In the following, we demonstrate that there is at most one singular interval, and further, we demonstrate that if there is a singular interval, then there is no switching point.

Consider that control $u^*(t)$ has a singular interval $[t_1, t_2]$, and, for simplicity, assume further $t_1 > 0$ and $t_2 < a$. The remaining cases, in which $t_1 = 0$ or $t_2 = a$, can be analyzed in

the same way. Figure 4.2(a) illustrates the possible trajectories for a singular interval $[t_1, t_2]$ with $0 < t_1 < t_2 < a$. According to Condition (4.24), we have $x^*(t) = 0$ for $t \in [t_1, t_2]$. According to the definition of a singular interval, we have $p^*(t) \neq 0$ in a neighborhood outside of the singular interval $[t_1, t_2]$, i.e., there exists an $\epsilon > 0$ such that $p^*(t) \neq 0$ for $t \in (t_1 - \epsilon, t_1) \cup (t_2, t_2 + \epsilon)$. The optimal control law according to Equation (4.17) thus requires $u^*(t) \neq d$ for $t \in (t_1 - \epsilon, t_1) \cup (t_2, t_2 + \epsilon)$. In particular, we have $u^*(t) = U$ for $p^*(t) < 0$, and $u^*(t) = 0$ for $p^*(t) > 0$. These production levels are represented in Figure 4.2(a) by segments (1) to (4). Segment (1) represents the case $u^*(t) = 0$ for $t \in (t_1 - \epsilon, t_1)$, segment (2) represents the case $u^*(t) = U$ for $t \in (t_1 - \epsilon, t_1)$, segment (3) represents the case $u^*(t) = U$ for $t \in (t_2, t_2 + \epsilon)$, and finally, segment (4) represents the case $u^*(t) = 0$ for $t \in (t_2, t_2 + \epsilon)$.

We now show that there is at most one singular interval within $[0, a]$. To the contrary, assume that there is another singular interval $[t_a, t_b]$ within $[0, a]$. By Condition (4.24), during the singular interval $[t_a, t_b]$, the production surplus level must be zero as well. A possible scenario is shown in Figure 4.2(b). From Figure 4.2(b), we see further that there must be a switching point with nonzero production surplus level between the two singular intervals. This contradicts Condition (4.24). Therefore, an optimal trajectory can have at most one singular interval.

We now show that it is not possible that an optimal control has a singular interval *and* a switching time $t_0 \in (0, a)$. Assume to the contrary that there is a switching point at time $t_0 \in (0, a)$ in addition to the singular interval $[t_1, t_2]$. By Condition (4.24), we must have $x(t_0) = 0$. In a similar way as above, we conclude that there must also be another switching point with nonzero production surplus level between the singular interval and the switching time t_0 . This again violates Condition (4.24). We conclude that an optimal control cannot have both a singular interval *and* a switching point.

Note that for $t_1 = 0$ or $t_2 = a$, we can apply similar reasoning to arrive at the same conclusion. Thus, if an optimal control has a singular interval $[t_1, t_2] \subseteq [0, a]$, it is necessary

that $[t_1, t_2]$ is the only singular interval and that no switching point exists. In summary, we have observed that until the trajectory reaches the zero production surplus level, the optimal control level is either U (if the initial production surplus is negative) or 0 (if the initial production surplus is positive). Subsequently, the singular interval starts. When the singular interval ends, the control will either be U or 0 . Thus, we conclude that it is necessary that the optimal control satisfies

$$u^*(t) = \begin{cases} \beta_1, & \text{if } t \in [0, t_1), \\ d, & \text{if } t \in [t_1, t_2), \\ \beta_2, & \text{if } t \in [t_2, a), \end{cases} \quad (4.25)$$

where β_1 and β_2 may take the value of 0 or U .

We now show that if the optimal control has a singular interval, another necessary condition for optimality is $\beta_2 = U$.

Before we proceed, we first establish a property which we will use to compare the performance of two policies, $u'(t)$ and $u''(t)$.

If two admissible controls $u'(t)$ and $u''(t)$ lead to trajectories $x'(t)$ and $x''(t)$

such that (a) $|x'(t)| \leq |x''(t)|$, for $t \in [0, a]$, and

(b) either $0 \geq x'(a) > x''(a)$,

or $db \leq x'(a) < x''(a)$,

then, we have $J_1(u') < J_1(u'')$. (4.26)

To see the validity of this property, notice that we have

$$J_1(u') = \int_0^a (x'(t))^2 dt + \int_0^b (x'(a) - dt)^2 dt \leq \int_0^a (x''(t))^2 dt + \int_0^b (x'(a) - dt)^2 dt,$$

since $|x'(t)| \leq |x''(t)|$, for $t \in [0, a]$. Further, both $0 \geq x'(a) > x''(a)$ and $db \leq x'(a) < x''(a)$ imply $(x'(a) - dt)^2 < (x''(a) - dt)^2$ for $t \in [0, b]$, and therefore, we have $J_1(u') < J_1(u'')$.

We now use Property (4.26) to demonstrate $\beta_2 = U$ in Equation (4.25) by showing that $\beta_2 = 0$ does not lead to an optimal control. Note that we only need to consider $t_2 < a$,

since the interval (t_2, a) vanishes for $t_2 = a$. In that case, β_2 is of no interest. For $t_2 < a$, Equation (4.25) with $\beta_2 = 0$ leads to a control $u_{\beta_2=0}(t)$ with

$$u_{\beta_2=0}(t) = \begin{cases} \beta_1, & \text{if } t \in [0, t_1), \\ d, & \text{if } t \in [t_1, t_2), \\ 0, & \text{if } t \in [t_2, a). \end{cases}$$

and a trajectory $x_{\beta_2=0}(t)$ with

$$x_{\beta_2=0}(t) = \begin{cases} (\beta_1 - d)(t - t_1), & \text{if } t \in [0, t_1], \\ 0, & \text{if } t \in [t_1, t_2], \\ -d(t - t_2), & \text{if } t \in [t_2, a]. \end{cases}$$

We now modify the production control $u_{\beta_2=0}(t)$ by letting $\beta_2 = d$. The resulting control policy is

$$u_{\beta_2=d}(t) = \begin{cases} \beta_1, & \text{if } t \in [0, t_1), \\ d, & \text{if } t \in [t_1, t_2), \\ d, & \text{if } t \in [t_2, a), \end{cases}$$

and leads to a trajectory $x_{\beta_2=d}(t)$ with

$$x_{\beta_2=d}(t) = \begin{cases} (\beta_1 - d)(t - t_1), & \text{if } t \in [0, t_1], \\ 0, & \text{if } t \in [t_1, t_2], \\ 0, & \text{if } t \in [t_2, a]. \end{cases}$$

Note that we have $|x_{\beta_2=d}(t)| \leq |x_{\beta_2=0}(t)|$ for $t \in [0, a]$, and $0 = x_{\beta_2=d}(a) > x_{\beta_2=0}(a)$. By Property (4.26), we therefore have $J_1(u_{\beta_2=d}) < J_1(u_{\beta_2=0})$. Thus, $u_{\beta_2=0}(t)$ is not optimal, and consequently, $\beta_2 = U$ in Equation (4.25) is a necessary condition for optimality.

In summary, we conclude that if $u^*(t)$ is a right-continuous optimal control which has a singular interval $[t_1, t_2] \subseteq [0, a]$, it is necessary that $u^*(t)$ satisfies

$$u^*(t) = \begin{cases} \beta_1, & \text{if } t \in [0, t_1), \\ d, & \text{if } t \in [t_1, t_2), \\ U, & \text{if } t \in [t_2, a), \end{cases} \quad (4.27)$$

where $\beta_1 \in \{0, U\}$.

Please note that an optimal production control which has a singular interval is completely specified by the parameters t_1 , t_2 , and β_1 .

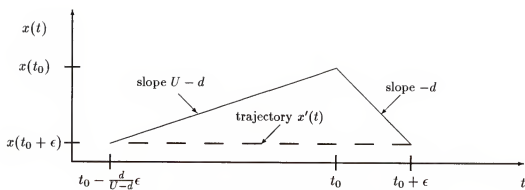
4.3.2 Switching Times and Switching Points

We now consider the case that there exists a *switching time* $t_0 \in (0, a)$. According to our observations above, we know that an optimal control cannot contain a singular interval and a switching point at the same time. In the following, we further show that there can be at most one switching point. In addition, we show that the switching point must satisfy $x^*(t_0) \geq 0$ and that the control must switch from 0 to U .

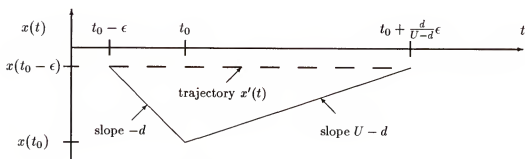
Recall that a switching time t_0 satisfies $p^*(t_0) = 0$ and represents the transition of $p^*(t)$ from a negative to a positive level or vice versa. According to Equation (4.17), we only need to consider a switch of control from U to 0 or from 0 to U . We further distinguish positivity and negativity of the production surplus level $x(t_0)$ at the switching time t_0 and show that switches from U to 0 and switches with $x(t_0) < 0$ are never optimal. Table 4.1 presents the cases for which we will show nonoptimality.

Switching from U to 0 at $t_0 \in (0, a)$ where $x(t_0) > 0$:

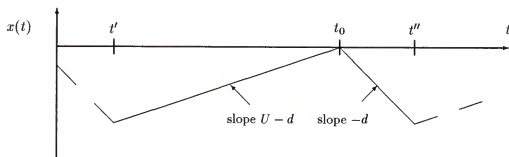
In the following, we show that a control with a switching point satisfying $x(t_0) > 0$, switching from U to 0 is not optimal by providing an improved control $u'(t)$. The idea is to truncate the trajectory around the switching point so that the modified trajectory stays closer to the zero production surplus level (see Figure 4.3(a)). In this case, there is a neighborhood around t_0 where the production control is U before t_0 and 0 after t_0 . Further,



(a) Switch from U to 0 at t_0 where $x(t_0) > 0$.



(b) Switch from 0 to U at t_0 where $x(t_0) < 0$.



(c) Switch from U to 0 at t_0 where $x(t_0) = 0$.

Figure 4.3. Switching Points which are not optimal.

the production surplus at t_0 is positive. Specifically, there exists an $\epsilon \in (0, \frac{x(t_0)}{d})$ such that $x(t) > 0$ for $t \in [t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon]$ and $u(t)$ satisfies

$$u(t) = \begin{cases} U, & \text{if } t \in [t_0 - \frac{d}{U-d}\epsilon, t_0), \\ 0, & \text{if } t \in [t_0, t_0 + \epsilon). \end{cases}$$

Notice that we have $x(t_0 - \frac{d}{U-d}\epsilon) = x(t_0 + \epsilon)$. Regardless of the production control $u(t)$ for $t \notin [t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon)$, we show that $u(t)$ is not optimal by improving the control during the interval $[t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon)$. The improved control $u'(t)$ is obtained as follows:

$$u'(t) = \begin{cases} u(t), & \text{if } t \notin [t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon), \\ d, & \text{if } t \in [t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon). \end{cases}$$

Consequently, we have

$$x'(t) = \begin{cases} x(t), & \text{if } t \notin [t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon], \\ x(t_0 + \epsilon), & \text{if } t \in [t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon]. \end{cases}$$

This implies that we have

$$J_1(u') = \int_0^a (x'(t))^2 dt + h_1(x'(a)) < \int_0^a (x(t))^2 dt + h_1(x'(a)),$$

since $x'^2(t) \leq x^2(t)$ for $t \in [0, a]$ and $x'^2(t) < x^2(t)$ for $t \in (t_0 - \frac{d}{U-d}\epsilon, t_0 + \epsilon)$. Since $x'(a) = x(a)$, we have $h_1(x'(a)) = h_1(x(a))$ and thus $J_1(u') < J_1(u)$. Therefore, a control which switches from U to 0 at a switching point with $x(t_0) > 0$ cannot be optimal.

Table 4.1. Classification of Switching Points.

Switching	$x(t_0) < 0$	$x(t_0) = 0$	$x(t_0) > 0$
from U to 0	not optimal	not optimal	not optimal
from 0 to U	not optimal	possibly optimal	possibly optimal

Switching from 0 to U at $t_0 \in (0, a)$ where $x(t_0) < 0$:

It can be shown that a control which switches from 0 to U at a switching point satisfying $x(t_0) < 0$ is not optimal either. The idea is to provide an improved control $u'(t)$ so that the modified trajectory $x'(t)$ is closer to the zero production level. This can be done in a similar way as above (see the dashed line in Figure 4.3(b)).

Switching from U to 0 at $t_0 \in (0, a)$ where $x(t_0) = 0$:

We show in two steps that a switch from U to 0 at time t_0 with switching point $x(t_0) = 0$ is not optimal. First, we show that there is no other switching time except t_0 . Second, we show that the production control which has no singular interval and which has exactly one switch at t_0 with $x(t_0) = 0$ is not optimal by providing an improved control.

Now assume that there is another switching time in addition to the switching time t_0 where the control switches from U to 0. Figure 4.3(c) illustrates this situation. The dashed line segments illustrate the situation that there is a switch at time t' (before t_0) and that there is a switch at time t'' (after t_0). Thus, if we have a second switching point either before t_0 or after t_0 , Figure 4.3(c) demonstrates that a second switching point would be a switching point which switches from 0 to U for negative production surplus. But we have just concluded above that a switch from 0 to U at negative production surplus level is not optimal. We therefore conclude that there can be no switching time other than t_0 . A control $u(t)$ which switches from U to 0 at $x(t_0) = 0$ has exactly one switching point,

$$u(t) = \begin{cases} U, & \text{if } t \in [0, t_0), \\ 0, & \text{if } t \in [t_0, a). \end{cases}$$

Next, we show that such a control is not optimal by providing an improved control $u'(t)$.

Consider the control $u'(t)$ with

$$u'(t) = \begin{cases} u(t), & \text{if } t \in [0, t_0), \\ d, & \text{if } t \in [t_0, a), \end{cases}$$

which leads to a trajectory $x'(t)$ with

$$x'(t) = \begin{cases} x(t), & \text{if } t \in [0, t_0], \\ 0, & \text{if } t \in [t_0, a]. \end{cases}$$

Thus, we have $|x'(t)| \leq |x(t)|$ for $t \in [0, a]$, and $0 = x'(a) > -d(a - t_0) = x(a)$. According to Property (4.26), we have $J_1(u') < J_1(u)$. In conclusion, a switching point at zero production surplus level which switches from U to 0 is not optimal.

Switching from U to 0 at $t_0 \in (0, a)$ where $x(t_0) < 0$:

In the same way in which we showed above that a switch from U to 0 at a zero production surplus level is not optimal, we can also show that a switch from U to 0 at a negative production surplus level is not optimal.

Summary

Thus far, we have shown that four of the six cases in Table 4.1 are not optimal. Therefore, an optimal control cannot have a switching point with negative production surplus, nor can a control which switches from U to 0 be optimal. Therefore, if an optimal control has a switching point, it is necessary that the control switches from 0 to U and that the production surplus at the switching time is nonnegative. We conclude that an optimal control can have at most one switching point, since a second switching point would have to switch from U to 0 . Further, nonnegative production surplus level at the switching time ($x^*(t_0) \geq 0$) requires positive initial production surplus level ($x(0) > 0$), since

$$x(0) = x^*(t_0) + dt_0 \geq dt_0 > 0, \quad \text{for } t_0 \in (0, a).$$

Thus, if $u^*(t)$ is an optimal control with a switching time $t_0 \in (0, a)$, necessary conditions for the switching point and the initial production surplus level are

$$x^*(t_0) \geq 0, \tag{4.28}$$

and

$$x(0) > 0, \tag{4.29}$$

Table 4.2. Initial production surpluses and initial control levels

	$u(0) = U$	$u(0) = d$	$u(0) = 0$
$x(0) < 0$		not optimal (Case A1)	not optimal (Case A2)
$x(0) = 0$	not optimal (Case A3)		not optimal (Case A4)
$x(0) > 0$	not optimal (Case A5)	not optimal (Case A6)	

and it is further necessary that $u^*(t)$ satisfies

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, t_0), \\ U, & \text{if } t \in [t_0, a). \end{cases} \quad (4.30)$$

4.3.3 Initial Control Level $u^*(0)$

We now examine the initial control level $u^*(0)$ in dependence of the initial production surplus level $x(0)$. We only need to consider the three control levels d , 0 , and U , since a singular interval requires $u^*(t) = d$, and if $p^*(t) \neq 0$, we have $u^*(t) = 0$ or $u^*(t) = U$ according to Equation (4.17). In the following, we will show $u^*(0) = 0$ for $x(0) > 0$, $u^*(0) = d$ for $x(0) = 0$, and $u^*(0) = U$ for $x(0) < 0$ by showing that all other cases are not optimal. The non-optimal cases are numbered in Table 4.2.

Case A1:

In this case, we have $u(0) = d$ for an initial production surplus $x(0) < 0$. Since a control level of d can only occur within a singular interval and since we have established that a singular interval can only occur at zero production surplus level (see Property (4.24)), $u(0) = d$ cannot be optimal for $x(0) < 0$.

Case A2:

In this case, we have $u(0) = 0$ for an initial production surplus $x(0) < 0$. We first note that such a control cannot have a switching point, since the initial production surplus is negative (compare Equation (4.29)). Thus, the production control $u(t)$ satisfies

$$u(t) = 0, \quad \text{for } t \in [0, a].$$

We define an alternative control $u_-(t)$ by

$$u_-(t) = d, \quad \text{for } t \in [0, a]. \quad (4.31)$$

The controls lead to trajectories $x(t)$ and $x_-(t)$, respectively, which satisfy

$$x(t) = x(0) - dt \leq x(0) < 0, \quad \text{for } t \in [0, a],$$

and

$$x_-(t) = x(0) < 0, \quad \text{for } t \in [0, a].$$

We conclude $|x(0)| = |x_-(t)| \leq |x(t)|$ for $t \in [0, a]$, and $0 > x_-(a) = x(0) > x(0) - da = x(a)$.

By Property (4.26), we have $J_1(u_-) < J_1(u)$ and thus, $u(0) = 0$ is not optimal for $x(0) < 0$.

Case A3:

In this case, we have $u(0) = U$ for an initial production surplus $x(0) = 0$. From the discussion in the previous section, such a control cannot have a switching point, since an optimal switch never switches from U to 0. Thus, the production control $u(t)$ satisfies

$$u(t) = U, \quad \text{for } t \in [0, a].$$

It leads to a trajectory with

$$x(t) = \begin{cases} (U - d)t, & \text{for } t \in [0, a], \\ (U - d)a - d(t - a), & \text{for } t \in [a, a + b]. \end{cases}$$

We now show that this trajectory is not optimal and distinguish two cases, $d < \frac{a}{a+b}U$ (Case A3-1) and $d = \frac{a}{a+b}U$ (Case A3-2).

Case A3-1:

In case of $d < \frac{a}{a+b}U$, we define an alternative control $u_+(t)$ by

$$u_+(t) = \begin{cases} d, & \text{if } t \in [0, a - \frac{db}{U-d}), \\ U, & \text{if } t \in [a - \frac{db}{U-d}, a]. \end{cases}$$

It leads to a trajectory with

$$x_+(t) = \begin{cases} 0, & \text{for } t \in [0, a - \frac{db}{U-d}], \\ (U-d)(t - (a - \frac{db}{U-d})), & \text{for } t \in [a - \frac{db}{U-d}, a], \\ db - d(t - a), & \text{for } t \in [a, a + b]. \end{cases}$$

Note that we have $x_+(t) \geq x_+(0) = 0$ for $t \in [0, a]$ and $x_+(a) = db$. Since $u_+(t) \leq u(t)$ for $t \in [0, a]$, we have $0 \leq x_+(t) \leq x(t)$ (and thus $|x_+(t)| \leq |x(t)|$) for $t \in [0, a]$, as well as $x(a) = x(0) + (U-d)a > db = x_+(a)$. According to Property (4.26), we have $J_1(u_+) < J_1(u)$.

Case A3-2:

In case of $d = \frac{a}{a+b}U$, we define a class of controls $u_{\Delta t}(t)$ by

$$u_{\Delta t}(t) = \begin{cases} d, & \text{if } t \in [0, \Delta t], \\ U, & \text{if } t \in [\Delta t, a], \end{cases}$$

where $\Delta t \geq 0$ is to be determined. It leads to a trajectory with

$$x_{\Delta t}(t) = \begin{cases} 0, & \text{for } t \in [0, \Delta t], \\ (U-d)(t - \Delta t), & \text{for } t \in [\Delta t, a], \\ (U-d)(a - \Delta t) - d(t - a), & \text{for } t \in [a, a + b]. \end{cases}$$

Notice that $x(t)$ is identical to $x_{\Delta t}(t)$ for $\Delta t = 0$. Since we have $x(a) = (U-d)a = db$ and $x_{\Delta t}(a) = db - (U-d)\Delta t < db$ for $\Delta t > 0$, we cannot apply Property (4.26). We directly use the definition of the performance measure and obtain

$$\begin{aligned} J_1(u_{\Delta t}) &= \int_{\Delta t}^a \{(U-d)(t - \Delta t)\}^2 dt + \int_a^{a+b} \{(U-d)(a - \Delta t) - d(t - a)\}^2 dt \\ &= \int_0^{a-\Delta t} \{(U-d)t\}^2 dt + \int_0^b \{(U-d)(a - \Delta t) - dt\}^2 dt. \end{aligned}$$

We take the derivative with respect to Δt and obtain

$$\begin{aligned}
\frac{d}{d\Delta t} J_1(u_{\Delta t}) &= (-1)\{(U-d)(a-\Delta t)\}^2 + \int_0^b -2(U-d)\{(U-d)(a-\Delta t) - dt\}dt \\
&= -(U-d)^2(a-\Delta t)^2 - 2(U-d)^2(a-\Delta t)b + (U-d)db^2 \\
&= -(U-d)^2\{(a-\Delta t)^2 + 2(a-\Delta t)b - \frac{d}{U-d}b^2\} \\
&= -(U-d)^2\{(a-\Delta t+b)^2 - \frac{U}{U-d}b^2\}.
\end{aligned} \tag{4.32}$$

From $d = \frac{a}{a+b}U$, we obtain $b = \frac{U-d}{d}a$. Replacing b in Equation (4.32) results in

$$\begin{aligned}
\frac{d}{d\Delta t} J_1(u_{\Delta t}) &= -(U-d)^2\left\{\left(\frac{U}{d}a - \Delta t\right)^2 - \frac{U}{U-d}\left\{\frac{U-d}{d}a\right\}^2\right\} \\
&= -(U-d)^2\left\{\frac{U^2}{d^2}\left(a - \frac{d}{U}\Delta t\right)^2 - \frac{U(U-d)}{d^2}\{a\}^2\right\} \\
&= -(U-d)^2\left\{\left(\frac{U^2}{d^2} - \frac{U(U-d)}{d^2}\right)a^2 - 2\frac{U}{d}a\Delta t + (\Delta t)^2\right\} \\
&= -(U-d)^2\left\{\frac{U}{d}a^2 - 2\frac{U}{d}a\Delta t + (\Delta t)^2\right\} \\
&= -(U-d)^2\left\{\frac{U}{d}a(a - 2\Delta t) + (\Delta t)^2\right\} \\
&< -(U-d)^2\left\{\frac{U}{d}a(a - 2\Delta t)\right\}.
\end{aligned}$$

That implies $\frac{d}{d\Delta t} J_1(u_{\Delta t}) < 0$ for $\Delta t < \frac{a}{2}$. Since $u(t)$ is identical to $u_{\Delta t}(t)$ for $\Delta t = 0$, we conclude that the policy $u(t)$ is not optimal. This concludes the considerations for Case A3-2.

Case A4:

In this case, we have $u(0) = 0$ for an initial production surplus $x(0) = 0$. In a similar way as in Case A2, we can show that $u_-(t)$ (Equation (4.31)) leads to a trajectory with improved performance, and thus, $u(0) = 0$ for $x(0) = 0$ is not optimal.

Case A5:

In this case, we have $u(0) = U$ for an initial production surplus $x(0) > 0$. For the same reason as in Case A3, we conclude that the optimal control $u(t)$ satisfies

$$u(t) = U, \quad \text{for } t \in [0, a),$$

which leads to a trajectory with

$$x(t) = \begin{cases} x(0) + (U - d)t, & \text{for } t \in [0, a], \\ x(0) + (U - d)a - d(t - a), & \text{for } t \in [a, a + b]. \end{cases}$$

We distinguish the two cases $x(0) \geq db$ (Case A5-1) and $x(0) \in (0, db)$ (Case A5-2).

Case A5-1:

For $x(0) \geq db$, we define an alternative control $u_d(t)$ by

$$u_d(t) = d, \quad \text{for } t \in [0, a].$$

It leads to a trajectory with

$$x_d(t) = \begin{cases} x(0), & \text{for } t \in [0, a], \\ x(0) - d(t - a), & \text{for } t \in [a, a + b]. \end{cases}$$

We have $|x_d(t)| \leq |x(t)|$ for $t \in [0, a]$ and

$$db \leq x_d(a) = x(0) < x(0) + (U - d)a = x(a).$$

Therefore, $J_1(u_d) < J_1(u)$ according to Property (4.26).

Case A5-2:

In this case, we have $x(0) \in (0, db)$. We define an alternative control $u_+(t)$ by

$$u_+(t) = \begin{cases} d, & \text{if } t \in [0, a - \frac{db - x(0)}{U - d}], \\ U, & \text{if } t \in [a - \frac{db - x(0)}{U - d}, a]. \end{cases}$$

Notice that $u_+(t)$ is well-defined, since

$$a - \frac{db - x(0)}{U - d} = \frac{a(U - d) - db + x(0)}{U - d} > 0,$$

and $a - \frac{db - x(0)}{U - d} < a$. The production rate $u_+(t)$ leads to a trajectory with

$$x_+(t) = \begin{cases} x(0), & \text{for } t \in [0, a - \frac{db - x(0)}{U - d}], \\ x(0) + (U - d)(t - (a - \frac{db - x(0)}{U - d})), & \text{for } t \in [a - \frac{db - x(0)}{U - d}, a], \\ db - d(t - a), & \text{for } t \in [a, a + b]. \end{cases}$$

Note that we have $x_+(t) \geq x_+(0) > 0$ for $t \in [0, a]$ and $x_+(a) = db$. Since $u_+(t) \leq u(t)$ for $t \in [0, a]$, we have $0 \leq x_+(t) \leq x(t)$ (and thus $|x_+(t)| \leq |x(t)|$) for $t \in [0, a]$, as well as $x(a) = x(0) + (U - d)a > db = x_+(a)$. According to Property (4.26), we have $J_1(u_+) < J_1(u)$.

Case A6:

In this case, we have $u(0) = d$ for an initial production surplus $x(0) > 0$. Since a control level of d can only occur within a singular interval and since we have established that a singular interval can only occur at zero production surplus level, $u(0) = d$ cannot be optimal for $x(0) > 0$.

In summary, we have for the initial control level $u^*(0)$:

$$u^*(0) = \begin{cases} 0, & \text{if } x(0) > 0, \\ d, & \text{if } x(0) = 0, \\ U, & \text{if } x(0) < 0. \end{cases} \quad (4.33)$$

Thus, for nonzero initial production surplus, the system initially moves toward zero production surplus level. If the production surplus is initially zero, we see that the optimal control will have a singular interval starting at 0.

4.3.4 Optimal Solution

In the previous subsections, we have seen that there are three possible structures for an optimal control. The optimal trajectory either has a singular interval or a switching point or neither of them. We have also seen that the initial control level $u^*(0)$ is determined by the initial production surplus $x(0)$ (Equation (4.33)). In this section, we solve for the optimal solution to the one-cycle problem P_1 with respect to the initial production surplus level $x(0)$. We first start with the case $x(0) = 0$. The optimal control for this case will provide a basis to proceed with the cases $x(0) < 0$ and $x(0) > 0$. The range of the initial production surplus is partitioned as follows:

- B1: $x(0) = 0$,
- B2: $x(0) \in (-(U - d)(a - C_\tau), 0)$,

$$\begin{aligned}
\text{B3:} \quad & x(0) \leq -(U-d)(a-C_r), \\
\text{B3-1:} \quad & x(0) \in [-(U-d)a, -(U-d)(a-C_r)], \\
\text{B3-2:} \quad & x(0) < -(U-d)a, \\
\text{B4:} \quad & x(0) \geq d(a-C_r), \\
\text{B4-1:} \quad & x(0) \geq d(a + \frac{b}{2}), \\
\text{B4-2:} \quad & x(0) \in [d(a-C_r), d(a + \frac{b}{2})], \\
\text{B5:} \quad & x(0) \in (0, d(a-C_r)),
\end{aligned}$$

where $C_r = b(\sqrt{\frac{U}{U-d}} - 1)$ is derived in Section 4.3.4.

In each case, we examine which of the three control structures is optimal. In order to conveniently refer to the structures, we number them as follows:

1. **Type-1 Structure:** The optimal trajectory has a singular interval $[t_1, t_2]$. According to Equation (4.27), we have for the optimal control

$$u^*(t) = \begin{cases} \beta_1, & \text{if } t \in [0, t_1), \\ d, & \text{if } t \in [t_1, t_2), \\ U, & \text{if } t \in [t_2, a), \end{cases}$$

where $\beta_1 \in \{0, U\}$.

2. **Type-2 Structure:** The optimal trajectory has a switching point at time t_0 . According to Equation (4.30), we have for the optimal control

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, t_0), \\ U, & \text{if } t \in [t_0, a). \end{cases}$$

Please note that we have in particular

$$u^*(0) = 0. \quad (4.34)$$

3. **Type-3 Structure:** The optimal trajectory has neither a singular interval nor a switching point. Since the costate function $p^*(t)$ is continuous, we thus have not only

$p^*(t) \neq 0$, but also

$$\begin{aligned} \text{either } p^*(t) < 0, & \quad \text{for } t \in [0, a], \\ \text{or } p^*(t) > 0, & \quad \text{for } t \in [0, a]. \end{aligned}$$

This implies for the optimal control

$$\begin{aligned} \text{either } u^*(t) &= 0, & \text{for } t \in [0, a], \\ \text{or } u^*(t) &= U, & \text{for } t \in [0, a]. \end{aligned} \tag{4.35}$$

Please note that this requires

$$u^*(0) \in \{0, U\}. \tag{4.36}$$

Case B1: $x(0) = 0$

According to Equation (4.33), we have $u^*(0) = d$. Due to Properties (4.34) and (4.36), this excludes a Type-2 Structure and a Type-3 Structure from consideration. Consequently, the optimal trajectory has a Type-1 Structure which has a singular interval $[t_1, t_2]$. Since $u^*(0) = d$, the singular interval has the left endpoint $t_1 = 0$. Substituting $t_1 = 0$ in Equation (4.27), we obtain

$$u(t) = \begin{cases} d, & \text{if } t \in [0, t_2], \\ U, & \text{if } t \in [t_2, a], \end{cases} \tag{4.37}$$

where $t_2 \in (0, a]$. In order to simplify the derivation, we introduce parameter C by defining $C = a - t_2$. The parameter C therefore satisfies $C \in [0, a)$. Substituting t_2 by $a - C$ in Equation (4.37) results in

$$u(t) = \begin{cases} d, & \text{if } t \in [0, a - C], \\ U, & \text{if } t \in [a - C, a], \end{cases} \tag{4.38}$$

which leads to the trajectory (see Figure 4.4),

$$x(t) = \begin{cases} 0, & \text{if } t \in [0, a - C], \\ (U - d)(t - (a - C)), & \text{if } t \in [a - C, a], \\ (U - d)C - d(t - a), & \text{if } t \in [a, a + b]. \end{cases}$$

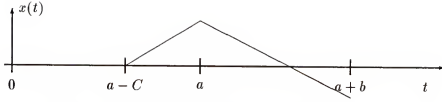


Figure 4.4. A Possible Trajectory for $x(0) = 0$.

We now seek to find the control with the best performance measure. Since both the optimal control and the trajectory are characterized by the parameter C , the performance measure can also be expressed in terms of C . We write $J_1(C)$ in place of $J_1(u)$,

$$J_1(C) = \int_0^a x^2(t)dt + h_1(x(a)) = \int_0^C ((U-d)t)^2 dt + \int_0^b ((U-d)C - dt)^2 dt. \quad (4.39)$$

We seek to determine the value of $C \in [0, a]$ which solves

$$\min\{J_1(C) : C \in [0, a]\}. \quad (4.40)$$

Therefore, we differentiate $J_1(C)$ by using Leibniz' rule

$$\frac{d}{dx} \int_{a(x)}^{b(x)} f(x, t) dt = f(x, b(x)) \frac{d}{dx} b(x) - f(x, a(x)) \frac{d}{dx} a(x) + \int_{a(x)}^{b(x)} \frac{d}{dx} f(x, t) dt,$$

and obtain

$$\begin{aligned} \frac{d}{dC} J_1(C) &= ((U-d)C)^2 + \int_0^b 2(U-d)((U-d)C - dt) dt \\ &= (U-d)^2 \{C^2 + 2bC - \frac{d}{U-d} b^2\}. \end{aligned} \quad (4.41)$$

The function $\frac{d}{dC} J_1(C)$ is of a quadratic form with roots

$$\begin{aligned} C'_r &= -b - \sqrt{b^2 + \frac{d}{U-d} b^2}, \\ C''_r &= -b + \sqrt{b^2 + \frac{d}{U-d} b^2}, \end{aligned} \quad (4.42)$$

where the smaller root is negative.

Since

$$\frac{d^2}{dC^2} J_1(C) = 2(U-d)^2(C+b),$$

we have

$$\begin{aligned}\frac{d^2}{dC^2} J_1(C'_r) &< 0, \\ \frac{d^2}{dC^2} J_1(C''_r) &> 0.\end{aligned}\tag{4.43}$$

Since we minimize, the unconstrained minimum is obtained for $C = C''_r$. We rename C''_r as C_r and have

$$C_r = -b + \sqrt{b^2 + \frac{d}{U-d}b^2} = b\left(\sqrt{\frac{U}{U-d}} - 1\right),$$

and

$$\frac{d^2}{dC^2} J_1(C_r) = (U-d)^2(2C_r + 2b) = (U-d)^2(2b\sqrt{\frac{U}{U-d}}) > 0.\tag{4.44}$$

It is not difficult to see that $C_r > 0$ since $\sqrt{\frac{U}{U-d}} > 1$. Further, using Condition (4.3) ($d \leq \frac{a}{a+b}U$), we also see that

$$\begin{aligned}C_r &= b\left(\sqrt{\frac{U}{U-d}} - 1\right) \leq b\left(\sqrt{\frac{U}{U-\frac{a}{a+b}U}} - 1\right) = b\left(\sqrt{\frac{1}{1-\frac{a}{a+b}}} - 1\right) \\ &= b\left(\sqrt{\frac{a+b}{b}} - 1\right) = \sqrt{b}\sqrt{a+b} - b < (\sqrt{a+b})^2 - b = a.\end{aligned}$$

Thus, C_r satisfies

$$C_r = b\left(\sqrt{\frac{U}{U-d}} - 1\right) \in (0, a).\tag{4.45}$$

Consequently, C_r solves the problem (4.40). This means that the singular interval is $[0, a - C_r]$ and the optimal control law satisfies

$$u^*(t) = \begin{cases} d, & \text{if } t \in [0, a - C_r], \\ U, & \text{if } t \in [a - C_r, a]. \end{cases}\tag{4.46}$$

Since we will frequently refer to the optimal trajectory for $x(0) = 0$, we name it $x_0^*(t)$. During the maintenance interval $(a, a + b)$, we have $u^*(t) = 0$ and together with Equation (4.46), we obtain for the optimal trajectory $x_0^*(t)$,

$$x_0^*(t) = \begin{cases} 0, & \text{if } t \in [0, a - C_r], \\ (U-d)(t - (a - C_r)), & \text{if } t \in [a - C_r, a], \\ (U-d)C_r - d(t - a), & \text{if } t \in [a, a + b]. \end{cases}\tag{4.47}$$

For the costate function $p^*(t)$, we have $p^*(t) = 0$ for $t \in [0, a - C_r]$. Using Equation (4.13), we have for $t \in [a - C_r, a]$,

$$\begin{aligned}
 p^*(t) &= p^*(t) - p^*(a - C_r) \\
 &= -2 \int_{a-C_r}^t x(\tau) d\tau \\
 &= -2 \int_{a-C_r}^t (U - d)(\tau - (a - C_r)) d\tau \\
 &= -2 \int_0^{t-(a-C_r)} (U - d)\tau d\tau \\
 &= -(U - d)(t - (a - C_r))^2.
 \end{aligned}$$

In summary, we have

$$p^*(t) = \begin{cases} 0, & \text{if } t \in [0, a - C_r], \\ -(U - d)(t - (a - C_r))^2, & \text{if } t \in [a - C_r, a]. \end{cases} \quad (4.48)$$

We verify the boundary condition (see Equation (4.15)):

$$\begin{aligned}
 2\{x^*(a)\}b - db^2 - \{p^*(a)\} &= 2\{(U - d)C_r\}b - db^2 - \{-(U - d)C_r^2\} \\
 &= (U - d)\{C_r^2 + 2bC_r - \frac{d}{U - d}b^2\} \\
 &= 0 \quad (\text{see Equation (4.41)}). \quad (4.49)
 \end{aligned}$$

The optimal control $u^*(t)$, the optimal trajectory $x^*(t)$, and the costate function $p^*(t)$ are given in Figure 4.5.

We will see that C_r is not only of significance in case of $x(0) = 0$. We will show that for all $x(0) \in (-(U - d)(a - C_r), d(a - C_r))$, the optimal control will result in a trajectory with a singular interval $[t_1, t_2]$ where the right endpoint of the singular interval is $t_2 = a - C_r$. A control which only assumes the minimum or the maximum production rate is called a bang-bang control. We will see that for an initial production surplus level with $x(0) \notin (-(U - d)(a - C_r), d(a - C_r))$, we will have a bang-bang control (with at most one switching point) as optimal policy. A bang-bang control with Type-2 Structure has one switching point and a bang-bang control with Type-3 Structure does not have a switching point.

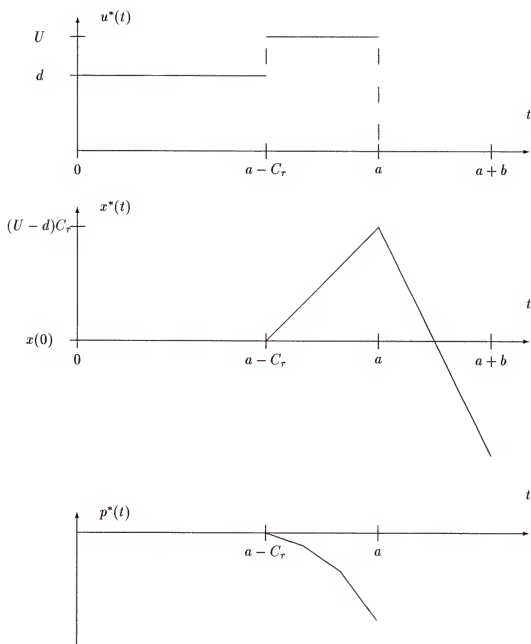


Figure 4.5. Optimal Solution for Case B1: $x(0) = 0$.

This concludes the analysis for the case where $x(0) = 0$. We now proceed with an optimality condition.

Optimality Condition

At this point, we will state an optimality condition for Problem P_1 which will be used in the reasoning of the remaining cases.

Optimality Condition

If $x_1^*(t)$ and $x_2^*(t)$ are two optimal trajectories for Problem P_1 with initial production surpluses $x_1^*(0)$ and $x_2^*(0)$, respectively, then the trajectories will not cross each other, i.e.,

$$\text{if } x_1^*(t') < x_2^*(t') \text{ for some } t' \in [0, a], \quad \text{then } x_1^*(t) \leq x_2^*(t) \text{ for all } t \in [0, a], \quad (4.50)$$

and in particular for $t' = 0$, we have

$$\text{if } x_1^*(0) < x_2^*(0), \quad \text{then} \quad x_1^*(t) \leq x_2^*(t), \quad \text{for } t \in [0, a]. \quad (4.51)$$

Explanation: We first note that the performance measure (see Equation (4.6)) is strictly convex, since

$$\frac{d^2}{dx^2(t)} \int_0^T x^2(t) dt = \int_0^T \frac{d^2}{dx^2(t)} x^2(t) dt = \int_0^T 2 \, dt = 2T > 0.$$

This implies that the optimal trajectory is unique for each initial production surplus. If the trajectories never meet, they don't cross each other either. And if the trajectories meet at a time t' , they will coincide from then on since the optimal trajectory is unique.

Case B2: $x(0) \in (-(U-d)(a-C_r), 0)$

First, since $x(0) < 0$, the initial control level satisfies $u^*(0) = U$ according to Equation (4.33). Next, we show that the optimal control has a Type-1 Structure:

A Type-2 Structure can be excluded, since it requires $u^*(0) = 0$ according to Property (4.34).

A Type-3 Structure according to Equation (4.35) would imply

$$u(t) = U, \quad \text{for } t \in [0, a],$$

$$x(a - C_r) = x(0) + (U - d)(a - C_r) > 0, \quad \text{for } x(0) \in (-(U - d)(a - C_r), 0).$$

Such a trajectory would cross trajectory $x_0^*(t)$ (see Equation (4.46)), since

$$\begin{aligned} x(0) &< 0 = x_0^*(0), \\ x(a - C_r) &> 0 = x_0^*(a - C_r). \end{aligned}$$

Therefore, we can also exclude a Type-3 Structure from consideration.

In conclusion, we must have a Type-1 Structure, i.e., the optimal trajectory has a singular interval. We now determine the left and the right endpoint of the singular interval $[t_1, t_2]$. The left endpoint t_1 is the time when the production surplus reaches the zero level, i.e., $t_1 = \frac{-x(0)}{U-d}$. At this point, the trajectory for $x(0) \in (-(U-d)(a-C_r), 0)$ meets with Trajectory $x_0^*(t)$. By the Optimality Condition (Equation (4.51)), the optimal trajectory for $x(0) \in (-(U-d)(a-C_r), 0)$ during the interval $[\frac{-x(0)}{U-d}, a]$ is identical to Trajectory $x_0^*(t)$. This means that the right endpoint of the singular interval is $t_2 = a - C_r$.

Therefore, we have for the optimal control $u^*(t)$,

$$u^*(t) = \begin{cases} U, & \text{if } t \in [0, \frac{-x(0)}{U-d}), \\ d, & \text{if } t \in [\frac{-x(0)}{U-d}, a - C_r), \\ U, & \text{if } t \in [a - C_r, a]. \end{cases} \quad (4.52)$$

The optimal trajectory $x^*(t)$ satisfies

$$x^*(t) = \begin{cases} (U-d)(t - \frac{-x(0)}{U-d}), & \text{if } t \in [0, \frac{-x(0)}{U-d}], \\ 0, & \text{if } t \in [\frac{-x(0)}{U-d}, a - C_r], \\ (U-d)(t - (a - C_r)), & \text{if } t \in [a - C_r, a], \\ (U-d)C_r - d(t - a), & \text{if } t \in [a, a + b]. \end{cases} \quad (4.53)$$

From the definition of a singular interval, we have $p^*(t) = 0$, for $t \in [\frac{-x(0)}{U-d}, a - C_r]$.

From Equation (4.13), we obtain for $t \in [0, \frac{-x(0)}{U-d}]$,

$$\begin{aligned} p^*(t) &= \int_t^{\frac{-x(0)}{U-d}} 2x^*(\tau) d\tau + p^*\left(\frac{-x(0)}{U-d}\right) \\ &= \int_t^{\frac{-x(0)}{U-d}} 2(U-d)(\tau - \frac{-x(0)}{U-d}) d\tau \\ &= -(U-d)\left(t - \frac{-x(0)}{U-d}\right)^2. \end{aligned}$$

Since the trajectory $x^*(t)$ for $t \in [\frac{-x(0)}{U-d}, a+b]$ is identical to the trajectory in Equation (4.47), the costate function $p^*(t)$ in the remaining interval $[\frac{-x(0)}{U-d}, a]$ yields the same expression for $p^*(t)$ as in Equation (4.48). In summary, we have

$$p^*(t) = \begin{cases} -(U-d)(t - \frac{-x(0)}{U-d})^2, & \text{if } t \in [0, \frac{-x(0)}{U-d}], \\ 0, & \text{if } t \in [\frac{-x(0)}{U-d}, a - C_r], \\ -(U-d)(t - (a - C_r))^2, & \text{if } t \in [a - C_r, a]. \end{cases} \quad (4.54)$$

Since the expressions for $x^*(a)$ and $p^*(a)$ are the same as in Case B1, the boundary condition (see Equation (4.15)) is verified in the same way as in Equation (4.49). The optimal control $u^*(t)$, the optimal trajectory $x^*(t)$, and the costate function $p^*(t)$ are shown in Figure 4.6.

Case B3: $x(0) < -(U-d)(a - C_r)$

First, since $x(0) < 0$, the initial control level satisfies $u^*(0) = U$ according to Equation (4.33). Next, we examine the control structure.

According to Property (4.34), we can exclude a Type-2 Structure (a production control with a switching point), since it requires $u^*(0) = 0$. Therefore, the optimal control has either a Type-1 Structure or a Type-3 Structure. That is, the optimal control either produces at maximum capacity or it has a singular interval. By distinguishing between the cases $x(0) \in [-(U-d)a, -(U-d)(a - C_r)]$ (Case B3-1) and $x(0) < -(U-d)a$ (Case B3-2), we will show that the optimal control has a Type-3 Structure.

Case B3-1: $x(0) \in [-(U-d)a, -(U-d)(a - C_r)]$.

In case of a Type-3 Structure, we produce at *maximum capacity* during $[0, a]$:

$$u(t) = U, \quad \text{for } t \in [0, a]. \quad (4.55)$$

In case of a Type-1 Structure, we have a *singular interval* $[t_1, t_2]$. The production control satisfies Equation (4.27) where $\beta_1 = u^*(0) = U$. The value of β_1 is determined according to the initial production control level (see Equation (4.33)). For the left endpoint of the

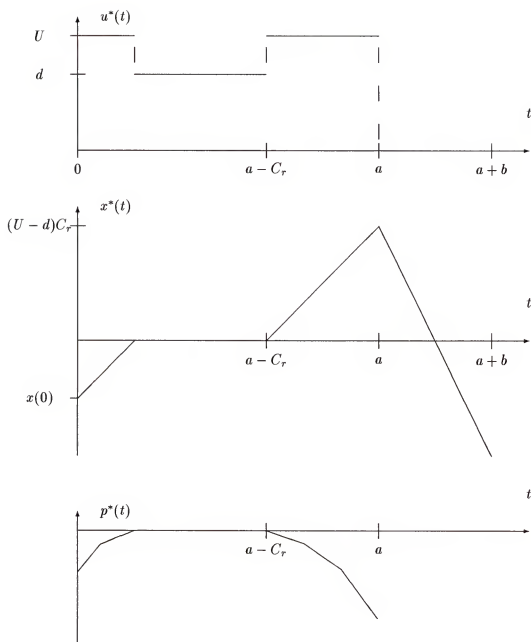


Figure 4.6. Optimal Solution for Case B2: $x(0) \in (-(U - d)(a - C_r), 0)$.

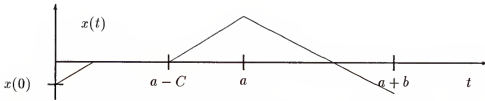


Figure 4.7. A possible trajectory for $x(0) \in [-(U-d)a, -(U-d)(a-C_r)]$.

singular interval, we have $t_1 = \frac{-x(0)}{U-d}$. Let $t_2 = a - C$ where $C \in [0, a - \frac{-x(0)}{U-d}]$. With $[t_1, t_2] = [\frac{-x(0)}{U-d}, a - C]$, the production control thus satisfies

$$u(t) = \begin{cases} U, & \text{if } t \in [0, \frac{-x(0)}{U-d}), \\ d, & \text{if } t \in [\frac{-x(0)}{U-d}, a - C], \\ U, & \text{if } t \in [a - C, a]. \end{cases} \quad (4.56)$$

Notice that we can represent the production control in Equation (4.55) by letting $C = a - \frac{-x(0)}{U-d}$ in Equation (4.56). Thus, if we extend the range of C from $[0, a - \frac{-x(0)}{U-d}]$ to $[0, a - \frac{-x(0)}{U-d}]$, we have represented both control structures which need to be considered.

A possible trajectory $x(t)$ is shown in Figure 4.7. The performance measure can be expressed as a function of the parameter C ,

$$J_1(C) = \int_0^{\frac{-x(0)}{U-d}} ((U-d)t)^2 dt + \int_0^C ((U-d)t)^2 dt + \int_0^b ((U-d)C - dt)^2 dt. \quad (4.57)$$

We seek to determine the value of C which solves

$$\min\{J_1(C) : C \in [0, a - \frac{-x(0)}{U-d}]\}. \quad (4.58)$$

Notice that the cost function $J_1(C)$ in Equation (4.57) differs from the cost function in Equation (4.39) only by the term $\int_0^{\frac{-x(0)}{U-d}} ((U-d)t)^2 dt$. This term is independent of C . Therefore, the derivatives of the cost functions in Equations (4.57) and (4.39) are identical (see Equation (4.41)):

$$\frac{d}{dC} J_1(C) = (U-d)^2 \{C^2 + 2bC - \frac{d}{U-d} b^2\}. \quad (4.59)$$

Since the function $\frac{d}{dC} J_1(C)$ is a quadratic function with negative function values between its two roots C'_r and $C_r = C''_r$ (see Equation (4.42)), i.e.,

$$\frac{d}{dC} J_1(C) < 0, \quad \text{for } C \in (C'_r, C''_r). \quad (4.60)$$

Since the smaller root is negative ($C'_r < 0$), and since

$$a - \frac{-x(0)}{U-d} \leq C_r = C''_r, \quad \text{for } x(0) \leq -(U-d)(a - C_r),$$

we have

$$[0, a - \frac{-x(0)}{U-d}) \subset (C'_r, C''_r). \quad (4.61)$$

From Equations (4.60) and (4.61), we conclude

$$\frac{d}{dC} J_1(C) < 0, \quad \text{for } C \in [0, a - \frac{-x(0)}{U-d}). \quad (4.62)$$

That is, if C increases within the interval $[0, a - \frac{-x(0)}{U-d})$, the function $J_1(C)$ decreases.

Consequently, $C = a - \frac{-x(0)}{U-d}$ solves (4.58), and Equation (4.56) with $C = a - \frac{-x(0)}{U-d}$ leads to

$$u^*(t) = U, \quad \text{for } t \in [0, a). \quad (4.63)$$

Case B3-2: $x(0) < -(U-d)a$.

Recall that we have already excluded a Type-2 Structure, since it requires $u^*(0) = 0$ for the initial control level. In this case, the optimal control cannot have a singular interval either (Type-1 Structure), since any possible trajectory would satisfy

$$x(t) \leq x(0) + (U-d)a < 0, \quad \text{for } t \in [0, a] \text{ and } x(0) < -(U-d)a.$$

Thus, we have a Type-3 Structure. Since the initial control level satisfies $u^*(0) = U$, we have the same optimal control as in Case B3-1 (see Equation (4.63)) and produce at full capacity.

We now consider the optimal trajectory and the costate function. Since we have the same optimal production control in both cases B3-1 and B3-2, we do not need to consider these two cases separately.

The optimal trajectory for $x(0) < -(U-d)(a - C_r)$ satisfies

$$x^*(t) = \begin{cases} x(0) + (U-d)t, & \text{if } t \in [0, a], \\ x(0) + (U-d)a - d(t-a), & \text{if } t \in [a, a+b]. \end{cases} \quad (4.64)$$

From Equation (4.13), we have $p^*(a) - p^*(t) = \int_t^a -2x^*(\tau)d\tau$, or equivalently,

$$\begin{aligned} p^*(t) &= 2 \int_t^a x^*(\tau)d\tau + p^*(a), \\ &= 2 \int_t^a (x(0) + (U-d)\tau)d\tau + p^*(a), \\ &= 2x(0)(a-t) + (U-d)(a^2 - t^2) + p^*(a), \quad \text{for } t \in [0, a]. \end{aligned} \quad (4.65)$$

From the boundary condition (Equation (4.15)), we obtain

$$p^*(a) = 2x^*(a)b - db^2 = 2\{x(0) + (U-d)a\}b - db^2. \quad (4.66)$$

From Equations (4.65) and (4.66), we obtain

$$p^*(t) = 2(a+b-t)x(0) + (U-d)(a^2 + 2ab - t^2) - db^2. \quad (4.67)$$

Notice that the costate function is a monotone increasing function since $p^*(t_2) - p^*(t_1) = \int_{t_1}^{t_2} -2x^*(\tau)d\tau \geq 0$ for $t_1 < t_2$ and that an upper bound for the costate function value is $-db^2$, since

$$p^*(a) = 2\{x(0) + (U-d)a\}b - db^2 < -db^2, \quad \text{for } x(0) < -(U-d)a.$$

These observations for the costate function $p^*(t)$ are reflected in Figure 4.8. It also illustrates the optimal control $u^*(t)$ and the optimal trajectory $x^*(t)$. This completes the analysis of Case B3-2 ($x(0) \leq -(U-d)a$).

In summary, with Cases B1, B2, and B3, we have now analyzed all cases which satisfy $x(0) \leq 0$. We have seen that the optimal control is to produce at *maximum capacity* for $x(0) \leq -(U-d)(a-C_r)$, and that there exists a *singular interval* for $x(0) \in (-(U-d)(a-C_r), 0]$. In the following, we will analyze Cases B4 and B5 which represent the two cases for which we have $x(0) > 0$.

Case B4: $x(0) > d(a-C_r)$.

First, since $x(0) > 0$, the initial control level satisfies $u^*(0) = 0$ according to Equation (4.33). Next, we can eliminate a Type-1 Structure from consideration. The reasoning is as follows:

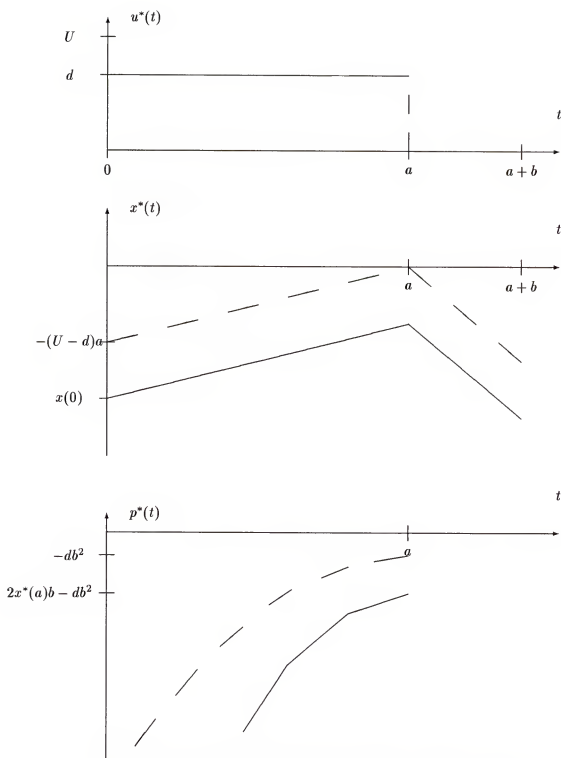


Figure 4.8. Optimal Solution for Case B3: $x(0) \leq -(U-d)(a-C_r)$.

For any production control, we have

$$x(t) \geq x(0) - dt > 0, \quad \text{for } t \in [0, a - C_r) \text{ and } x(0) \geq d(a - C_r). \quad (4.68)$$

Since an optimal trajectory $x^*(t)$ for $x(0) \geq d(a - C_r)$ does not cross the Trajectory $x_0^*(t)$ (see Equation (4.51)), and since the Trajectory $x_0^*(t)$ satisfies $x_0^*(t) > 0$ for $t \in (a - C_r, a]$, we have

$$x^*(t) \geq x_0^*(t) > 0, \quad \text{for } t \in (a - C_r, a] \text{ and } x(0) \geq d(a - C_r). \quad (4.69)$$

From Equations (4.68) and (4.69), there cannot be a singular interval. We therefore exclude a Type-1 Structure from consideration.

Thus, we have either a Type-2 Structure or a Type-3 Structure: Since we have $u^*(0) = 0$ for $x(0) > 0$ according to Equation (4.33), we either *shut* the machine *down* during the entire interval or we switch the production rate from 0 to U at some time $t_0 \in (0, a)$.

If we *shut* the machine *down*, the control policy is

$$u(t) = 0, \quad \text{for } t \in [0, a). \quad (4.70)$$

If we have a policy with a *switching point* at time $t_0 \in (0, a)$, we introduce a parameter S with $S = a - t_0 \in (0, a)$ to simplify the calculation. According to Equation (4.30) with switching time $t_0 = a - S$, we have

$$u(t) = \begin{cases} 0, & \text{if } t \in [0, a - S), \\ U, & \text{if } t \in [a - S, a). \end{cases} \quad (4.71)$$

We now combine the solution structures. Note that if we let $S = 0$ in Equation (4.71), the interval $[a - S, a]$ vanishes and we obtain the same policy as in Equation (4.70).

Thus, the optimal control for $x(0) \geq d(a - C_r)$ is specified by the optimal value of S where $S \in [0, a)$. The trajectory is

$$x(t) = \begin{cases} x(0) - dt, & \text{if } t \in [0, a - S), \\ x(0) - d(a - S) + (U - d)(t - (a - S)), & \text{if } t \in [a - S, a). \end{cases} \quad (4.72)$$

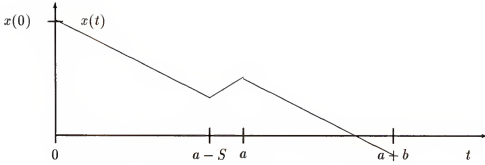


Figure 4.9. A Possible Trajectory for $x(0) \geq d(a - C_r)$.

It is shown in Figure 4.9. The performance measure $J_1(S)$ satisfies

$$\begin{aligned} J_1(S) = & \int_0^{a-S} (x(0) - dt)^2 dt + \int_0^S (x(0) - d(a - S) + (U - d)t)^2 dt \\ & + \int_0^b (x(0) - da + US - dt)^2 dt. \end{aligned} \quad (4.73)$$

Thus, we are seeking the value of parameter S which solves

$$\min\{J_1(S) : S \in [0, a]\}. \quad (4.74)$$

For the derivative, we have according to Leibniz' rule

$$\begin{aligned} \frac{d}{dS} J_1(S) = & -1(x(0) - d(a - S))^2 + (x(0) - d(a - S) + (U - d)S)^2 \\ & + \int_0^S 2d(x(0) - d(a - S) + (U - d)t)dt + \int_0^b 2U(x(0) - da + US - dt)dt. \end{aligned}$$

Evaluating the integrals and rearranging the terms, we have

$$\frac{d}{dS} J_1(S) = U(U + d)\{S^2 + 2\frac{x(0) - da + Ub}{U + d}S - \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))\}. \quad (4.75)$$

The two roots S'_1 and S''_1 are given by

$$\begin{aligned} S'_1 &= -\frac{x(0) - da + Ub}{U + d} - \sqrt{\left(\frac{x(0) - da + Ub}{U + d}\right)^2 + \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))}, \\ S''_1 &= -\frac{x(0) - da + Ub}{U + d} + \sqrt{\left(\frac{x(0) - da + Ub}{U + d}\right)^2 + \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))}, \end{aligned}$$

and satisfy

$$\begin{aligned}\frac{d^2}{dS^2}J_1(S'_1) &= -2U(U+d)\sqrt{\left(\frac{x(0)-da+Ub}{U+d}\right)^2 + \frac{2b}{U+d}\left(d\left(a+\frac{b}{2}\right)-x(0)\right)}, \\ \frac{d^2}{dS^2}J_1(S''_1) &= 2U(U+d)\sqrt{\left(\frac{x(0)-da+Ub}{U+d}\right)^2 + \frac{2b}{U+d}\left(d\left(a+\frac{b}{2}\right)-x(0)\right)}.\end{aligned}\quad (4.76)$$

The smaller root S'_1 is negative and satisfies $\frac{d^2}{dS^2}J_1(S'_1) < 0$. The larger root is S''_1 and satisfies $\frac{d^2}{dS^2}J_1(S''_1) > 0$. Since our objective is to minimize, we focus on S''_1 . We rename S''_1 as $S_1(x(0))$:

$$S_1(x(0)) = -\frac{x(0)-da+Ub}{U+d} + \sqrt{\left(\frac{x(0)-da+Ub}{U+d}\right)^2 + \frac{2b}{U+d}\left(d\left(a+\frac{b}{2}\right)-x(0)\right)}.\quad (4.77)$$

According to Equation (4.76), we have

$$\frac{d^2}{dS^2}J_1(S_1(x(0))) > 0.\quad (4.78)$$

At this point, we observe in Equation (4.75) that $\frac{d}{dS}J_1(S)$ is negative between its roots,

$$\frac{d}{dS}J_1(S) < 0, \quad \text{for } S \in (S'_1, S''_1).$$

and positive for $S > S_1(x(0)) = S''_1$, i.e.,

$$\frac{d}{dS}J_1(S) > 0, \quad \text{for } S > S_1(x(0)).\quad (4.79)$$

In the following, we consider the two cases $S_1(x(0)) \leq 0$ (Case B4-1) and $S_1(x(0)) > 0$ (Case B4-2). We will see that the optimal control for Case B4-1 has a Type-3 Structure while the optimal control for Case B4-2 has a Type-2 Structure.

Notice the following relationship between $x(0)$ and $S_1(x(0))$. Since we have $\frac{2b}{U+d}\left(d\left(a+\frac{b}{2}\right)-x(0)\right) \leq 0$, if and only if $x(0) \geq d\left(a+\frac{b}{2}\right)$, we conclude from Equation (4.77)

$$S_1(x(0)) \leq 0, \quad \text{iff } x(0) \geq d\left(a+\frac{b}{2}\right).\quad (4.80)$$

Thus, Case B4-1 is represented by $x(0) \geq d\left(a+\frac{b}{2}\right)$.

Further, Equation (4.80) and $S_1(x(0)) > 0$ imply $x(0) < d\left(a+\frac{b}{2}\right)$. Since Case B4 considers an initial surplus $x(0)$ satisfying $x(0) \geq d(a-C_r)$, Case B4-2 therefore covers

$$x(0) \in [d(a-C_r), d\left(a+\frac{b}{2}\right)).$$

Case B4-1: $x(0) \geq d(a + \frac{b}{2})$.

In this case, we will show that $S = 0$ solves (4.74). We first consider the derivative of the performance measure for $S \neq 0$. Since $x(0) \geq d(a + \frac{b}{2})$ implies $S_1(x(0)) \leq 0$ and $S_1(x(0)) \leq 0$ and $S \in (0, a)$ imply $S > S_1(x(0))$, we have according to Equations (4.79) and (4.80)

$$\frac{d}{dS} J_1(S) > 0, \quad \text{for } S \in (0, a) \text{ and } x(0) \geq d(a + \frac{b}{2}).$$

That is, as S increases, so does $J_1(S)$. Consequently, $S = 0$ solves (4.74). Substituting $S = 0$ in Equation (4.71) leads to the control

$$u^*(t) = 0, \quad \text{for } t \in [0, a], \quad (4.81)$$

i.e., there is neither a switching point nor a singular interval. Figure 4.11 illustrates this situation. The optimal trajectory $x^*(t)$ therefore satisfies

$$x^*(t) = x(0) - dt, \quad \text{for } t \in [0, a + b]. \quad (4.82)$$

From the boundary condition (see Equation (4.15)), we obtain $p^*(a) = 2x^*(a)b - db^2$.

Together with Equation (4.13), this leads to

$$\begin{aligned} p^*(t) &= 2 \int_t^a x^*(\tau) d\tau + p^*(a) \\ &= 2 \int_t^a (x(0) - d\tau) d\tau + 2(x(0) - da)b - db^2 \\ &= 2x(0)(a - t) - d(a^2 - t^2) + 2(x(0) - da)b - db^2 \\ &= 2x(0)(a + b - t) - d((a + b)^2 - t^2), \quad \text{for } t \in [0, a]. \end{aligned} \quad (4.83)$$

This completes our analysis for an initial production surplus satisfying $x(0) \geq d(a + \frac{b}{2})$ which corresponds to $S_1(x(0)) \leq 0$. The functions $u^*(t)$, $x^*(t)$, and $p^*(t)$ are shown in Figure 4.10.

Case B4-2: $x(0) \in [d(a - C_r), d(a + \frac{b}{2})]$.

Recall that the range for parameter S in (4.74) is $[0, a]$. In Appendix C.1 (see Equation (C.1)), we show

$$S_1(x(0)) \leq C_r < a, \quad \text{for } x(0) \in [d(a - C_r), d(a + \frac{b}{2})].$$

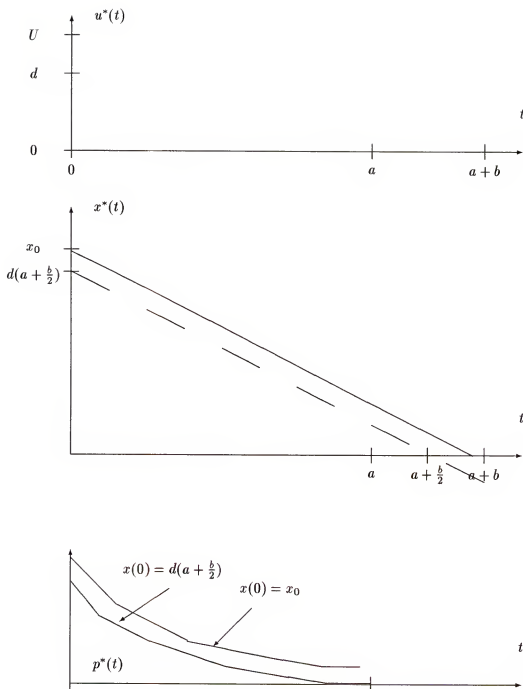


Figure 4.10. Optimal Solution for Case B4-1: $x(0) \geq d(a + \frac{b}{2})$.

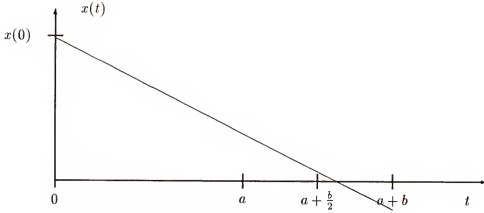


Figure 4.11. A possible trajectory for $x(0) \geq d(a + \frac{b}{2})$

Since we already have $S_1(x(0)) > 0$, this results in

$$S_1(x(0)) \in (0, C_r] \subseteq [0, a), \quad \text{for } x(0) \in [d(a - C_r), d(a + \frac{b}{2})]. \quad (4.84)$$

Since the second derivative of $J_1(S)$ for $S = S_1(x(0))$ is negative according to Equation (4.78), we conclude that $S_1(x(0))$ solves (4.74) and that the optimal control has a Type-2 Structure with a switching point at time $t_0 = a - S_1(x(0))$.

However, in order for the optimal control to have a switching point, the production surplus at the switching time must be nonnegative. In Appendix C.2, we verify that the optimality condition (see Equation (4.28)) is satisfied,

$$x^*(a - S_1(x(0))) \geq 0, \quad \text{for } x(0) \in [d(a - C_r), d(a + \frac{b}{2})].$$

The optimal policy is thus

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, a - S_1(x(0))], \\ U, & \text{if } t \in [a - S_1(x(0)), a]. \end{cases} \quad (4.85)$$

For the trajectory $x^*(t)$, we have

$$x^*(t) = \begin{cases} x(0) - dt, & \text{if } t \in [0, a - S_1(x(0))], \\ x(0) - dt + U(t - (a - S_1(x(0))))], & \text{if } t \in [a - S_1(x(0)), a], \\ x(0) - dt + U S_1(x(0)), & \text{if } t \in [a, a + b]. \end{cases} \quad (4.86)$$

For the function $p^*(t)$, we have a root at the switching point, i.e., $p^*(a - S_1(x(0))) = 0$.

Together with Equation (4.13), we derive

$$p^*(t) = \begin{cases} \int_t^{a-S_1(x(0))} 2x^*(\tau) d\tau, & \text{if } t \in [0, a - S_1(x(0))], \\ \int_{a-S_1(x(0))}^t -2x^*(\tau) d\tau, & \text{if } t \in [a - S_1(x(0)), a]. \end{cases}$$

For $t \in [0, a - S_1(x(0))]$, this yields

$$\begin{aligned} p^*(t) &= \int_t^{a-S_1(x(0))} 2x^*(\tau) d\tau = \int_t^{a-S_1(x(0))} 2(x(0) - d\tau) d\tau \\ &= 2x(0)(a - S_1(x(0)) - t) - d((a - S_1(x(0)))^2 - t^2). \end{aligned}$$

For $t \in [a - S_1(x(0)), a]$, this yields

$$\begin{aligned} p^*(t) &= \int_{a-S_1(x(0))}^t -2x^*(\tau) d\tau \\ &= -2 \int_{a-S_1(x(0))}^t \{x(0) - d(a - S_1(x(0))) + (U - d)(\tau - (a - S_1(x(0))))\} d\tau \\ &= -2 \int_{a-S_1(x(0))}^t \{x(0) - U(a - S_1(x(0))) + (U - d)\tau\} d\tau \\ &= -2(x(0) - U(a - S_1(x(0))))(t - (a - S_1(x(0)))) \\ &\quad - (U - d)(t^2 - (a - S_1(x(0)))^2). \end{aligned}$$

In summary, we have

$$p^*(t) = \begin{cases} 2x(0)(a - S_1(x(0)) - t) - d((a - S_1(x(0)))^2 - t^2), \\ \quad \text{if } t \in [0, a - S_1(x(0))]; \\ -2(x(0) - U(a - S_1(x(0))))(t - (a - S_1(x(0)))) \\ \quad - (U - d)(t^2 - (a - S_1(x(0)))^2), \\ \quad \text{if } t \in [a - S_1(x(0)), a]. \end{cases} \quad (4.87)$$

Finally, we verify the boundary condition (4.15). With $x^*(a) = x(0) - da + US_1(x(0))$

and

$$\begin{aligned} p^*(a) &= -2(x(0) - U(a - S_1(x(0))))(a - (a - S_1(x(0)))) \\ &\quad - (U - d)(a^2 - (a - S_1(x(0)))^2) \end{aligned}$$

$$\begin{aligned}
&= -2(x(0) - U(a - S_1(x(0))))S_1(x(0)) - (U - d)(2aS_1(x(0)) - S_1(x(0))^2) \\
&= -(U + d)(S_1(x(0)))^2 - 2(x(0) - da)S_1(x(0)),
\end{aligned}$$

we have

$$\begin{aligned}
&2x^*(a)b - db^2 - p^*(a) \\
&= 2\{x(0) - da + US_1(x(0))\}b - db^2 + (U + d)(S_1(x(0)))^2 + 2(x(0) - da)S_1(x(0)) \\
&= (U + d)(S_1(x(0)))^2 + 2(x(0) - da + Ub)S_1(x(0)) + 2(x(0) - d(a + \frac{b}{2}))b \\
&= (U + d)\{(S_1(x(0)))^2 + 2\frac{x(0) - da + Ub}{U + d}S_1(x(0)) + 2\frac{b}{U + d}(x(0) - d(a + \frac{b}{2}))\}(4.88) \\
&= 0. \tag{4.89}
\end{aligned}$$

Equation (4.89) is established by observing that $S_1(x(0))$ is a root of $\frac{d}{dS}J_1(S)$ (see Equation (4.75)) and that $\frac{d}{dS}J_1(S_1(x(0)))$ is a multiple of the term on the right-hand-side of Equation (4.88).

The functions $u^*(t)$, $x^*(t)$, and $p^*(t)$ are shown in Figure 4.12. This completes the analysis for Case B4-2 ($x(0) \in [d(a - C_r), d(a + \frac{b}{2})]$) which corresponds to $S_1(x(0)) > 0$ and $x(0) \geq d(a - C_r)$.

Case B5: $x(0) \in (0, d(a - C_r))$

First, since $x(0) > 0$, the initial control level satisfies $u^*(0) = 0$ according to Equation (4.33). We will see that the optimal control structure is a Type-1 Structure. We eliminate a Type-2 Structure and a Type-3 Structure according to the following reasoning:

According to the optimality condition (see Equation (4.51)), the optimal trajectory $x^*(t)$ for $x(0) \in (0, d(a - C_r))$ cannot cross the Trajectory $x_0^*(t)$, i.e.,

$$x^*(t) \geq x_0^*(t) \geq 0, \quad \text{for } t \in [0, a]. \tag{4.90}$$

With $u^*(0) = 0$, a Type-3 Structure would require $u(t) = 0$ for $t \in [0, a]$ and lead to

$$x(a - C_r) = x(0) - d(a - C_r) < 0, \quad \text{for } x(0) \in (0, d(a - C_r)).$$

Therefore, such a control cannot be optimal.

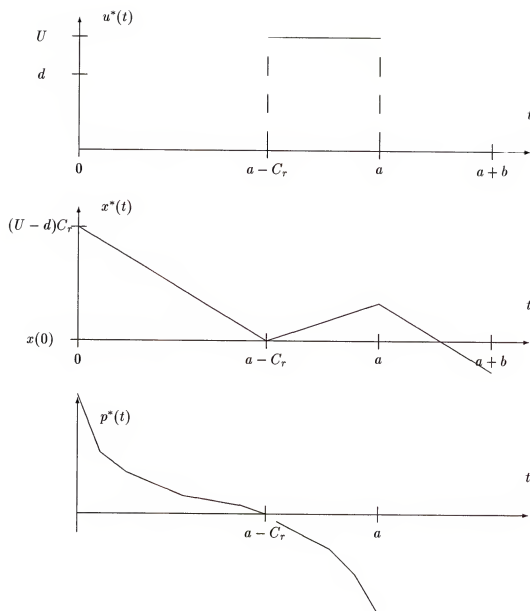


Figure 4.12. Optimal Solution for Case B4-2: $x(0) \in [d(a - C_r), d(a + \frac{b}{2})]$.

In order to show that a Type-2 Structure is not optimal, we consider the optimal trajectory for $x(0) = d(a - C_r)$ and in particular, we are interested in the production surplus level at time $t = a - C_r$. According to Appendix C.1 (see Equation (C.6)), we have

$$S_1(d(a - C_r)) = C_r.$$

Thus, we conclude that the switching time for $x(0) = d(a - C_r)$ is

$$a - S_1(d(a - C_r)) = a - C_r.$$

This implies that the optimal production surplus level for $x(0) = d(a - C_r)$ at time $t = a - C_r$ is zero:

$$x^*(a - C_r) = \{d(a - C_r)\} - d(a - \{S_1(d(a - C_r))\}) = 0. \quad (4.91)$$

We can now apply the optimality condition in Equation (4.51) and conclude that the production surplus level for an optimal trajectory with $x(0) \in (0, d(a - C_r))$ at time $t = a - C_r$ must be nonpositive according to Equation (4.91):

$$x^*(a - C_r) \leq 0, \quad \text{for } x(0) \in (0, d(a - C_r)). \quad (4.92)$$

From Equations (4.90) and (4.92), we can now conclude that a Type-2 Structure is not optimal: Let $u(t)$ be a production policy with switching time t_0 . According to Equation (4.90), we conclude $t_0 \leq \frac{x(0)}{d}$ ($< a - C_r$) and $x(t_0) \geq 0$. This implies

$$x(a - C_r) = x(0) - dt_0 + (U - d)(a - C_r - t_0) > x(0) - dt_0 = x(t_0) \geq 0. \quad (4.93)$$

Since Equation (4.93) violates Equation (4.92), we cannot have a Type-2 Structure.

In conclusion, we must have a Type-1 Structure, i.e., the optimal trajectory has a *singular interval*. The left endpoint satisfies $t_1 = \frac{x(0)}{d}$ at which the trajectory $x^*(t)$ meets with Trajectory $x_0^*(t)$. According to the Optimality Condition (see Equation (4.51)), the trajectories $x^*(t)$ and $x_0^*(t)$ will be identical from then on. That is, the right endpoint of the singular interval is $t_2 = a - C_r$.

Thus, the optimal control has the singular interval $[\frac{x(0)}{d}, a - C_r]$ and satisfies

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, \frac{x(0)}{d}], \\ d, & \text{if } t \in [\frac{x(0)}{d}, a - C_r], \\ U, & \text{if } t \in [a - C_r, a]. \end{cases} \quad (4.94)$$

The optimal trajectory $x^*(t)$ satisfies

$$x^*(t) = \begin{cases} d(\frac{x(0)}{d} - t), & \text{if } t \in [0, \frac{x(0)}{d}], \\ 0, & \text{if } t \in [\frac{x(0)}{d}, a - C_r], \\ (U - d)(t - (a - C_r)), & \text{if } t \in [a - C_r, a], \\ (U - d)C_r - d(t - a), & \text{if } t \in [a, a + b]. \end{cases} \quad (4.95)$$

Concerning the function $p^*(t)$, we must have $p^*(\frac{x(0)}{d}) = 0$ and $p^*(\frac{x(0)}{d}) - p^*(t) = -p^*(t) = \int_t^{\frac{x(0)}{d}} -2x^*(\tau)d\tau$. Therefore, we have for $t \in [0, \frac{x(0)}{d}]$,

$$p^*(t) = \int_t^{\frac{x(0)}{d}} 2x^*(\tau)d\tau = \int_t^{\frac{x(0)}{d}} 2d(\frac{x(0)}{d} - \tau)d\tau = -d(\frac{x(0)}{d} - t)^2 \Big|_t^{\frac{x(0)}{d}} = d(\frac{x(0)}{d} - t)^2.$$

The remaining interval $[\frac{x(0)}{d}, a]$ results in the same expression for $p^*(t)$ as in Cases B1 and B2 (see Equations (4.48) and (4.54)), since the trajectory $x^*(t)$ for $t \in [a - C_r, a + b]$ is identical to the optimal trajectories in Cases B1 and B2. In summary, we have

$$p^*(t) = \begin{cases} d(\frac{x(0)}{d} - t)^2, & \text{if } t \in [0, \frac{x(0)}{d}], \\ 0, & \text{if } t \in [\frac{x(0)}{d}, a - C_r], \\ -(U - d)(t - (a - C_r))^2, & \text{if } t \in [a - C_r, a]. \end{cases} \quad (4.96)$$

The boundary condition is verified in the same way as in Cases B1 and B2, since the expressions for $x^*(a)$ and $p^*(a)$ are identical to those for the Cases B1 and B2. The functions $u^*(t)$, $x^*(t)$, and $p^*(t)$ are illustrated in Figure 4.13. This completes our discussion of the case $x(0) \in (0, d(a - C_r))$.

Characterization of the optimal solution

We now summarize Equations (4.46), (4.52), (4.63), (4.81), (4.85), and (4.94). The optimal control policy $u_1^*(t)$ for the one-cycle problem P_1 is governed by

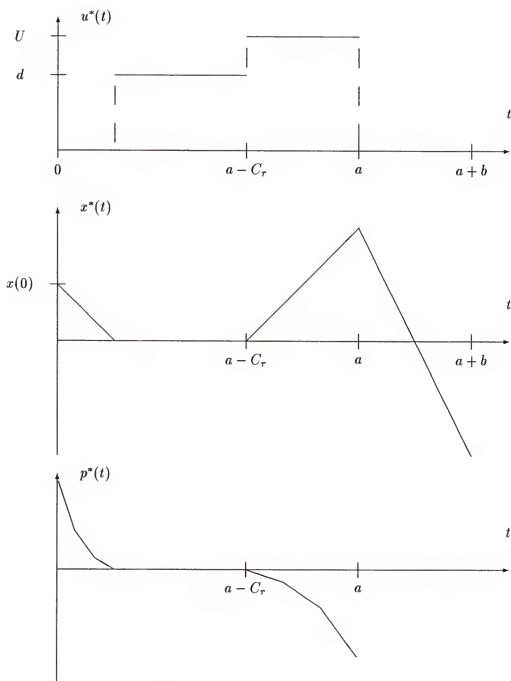
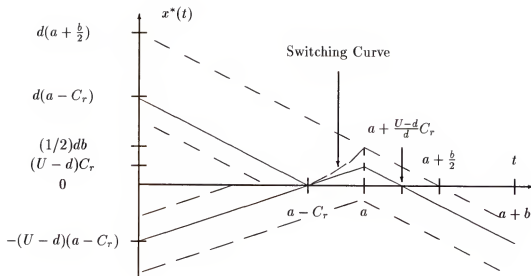


Figure 4.13. Optimal Solution for Case B5: $x(0) \in (0, d(a - C_r))$.

$$u_1^*(t) = \left\{ \begin{array}{ll} U, & \text{if } x(0) \leq -(U-d)(a-C_r), \quad \text{and} \quad t \in [0, a); \\ \dots\dots\dots \\ U, & \text{if } x(0) \in (-(U-d)(a-C_r), 0), \quad \text{and} \quad t \in [0, \frac{-x(0)}{U-d}), \\ d, & \text{if } x(0) \in (-(U-d)(a-C_r), 0), \quad \text{and} \quad t \in [\frac{-x(0)}{U-d}, a-C_r), \\ U, & \text{if } x(0) \in (-(U-d)(a-C_r), 0), \quad \text{and} \quad t \in [a-C_r, a); \\ \dots\dots\dots \\ d, & \text{if } x(0) = 0, \quad \text{and} \quad t \in [0, a-C_r), \\ U, & \text{if } x(0) = 0, \quad \text{and} \quad t \in [a-C_r, a); \\ \dots\dots\dots \\ 0, & \text{if } x(0) \in (0, d(a-C_r)), \quad \text{and} \quad t \in [0, \frac{x(0)}{d}), \\ d, & \text{if } x(0) \in (0, d(a-C_r)), \quad \text{and} \quad t \in [\frac{x(0)}{d}, a-C_r), \\ U, & \text{if } x(0) \in (0, d(a-C_r)), \quad \text{and} \quad t \in [a-C_r, a); \\ \dots\dots\dots \\ 0, & \text{if } x(0) \in [d(a-C_r), d(a+\frac{b}{2})], \quad \text{and} \quad t \in [0, a-S_1(x(0))], \\ U, & \text{if } x(0) \in [d(a-C_r), d(a+\frac{b}{2})], \quad \text{and} \quad t \in [a-S_1(x(0)), a); \\ \dots\dots\dots \\ 0, & \text{if } x(0) \geq d(a+\frac{b}{2}), \quad \text{and} \quad t \in [0, a). \end{array} \right. \quad (4.97)$$

We observe that we have a *singular interval* for $x(0) \in (-(U-d)(a-C_r), d(a-C_r))$, a *switching point* for $x(0) \in [d(a-C_r), d(a+\frac{b}{2})]$, maximum production during $[0, a]$ for $x(0) \leq -(U-d)(a-C_r)$, and no production for $x(0) \geq d(a+\frac{b}{2})$.

Figure 4.14 depicts some optimal trajectories: For an initial production surplus level $x(0)$ within the range $[-(U-d)(a-C_r), d(a-C_r)]$, the trajectory will reach the zero production surplus level at a time $t \in [0, a-C_r]$. The trajectory will then stay at the zero

Figure 4.14. Optimal Trajectory for Problem P_1 .

production surplus level until time $t = a - C_r$. For $t \in [a - C_r, a]$, the machine is operated at capacity before the maintenance takes place. The machine is shut down during the interval $[a, a + b]$.

Since all optimal controls with singular interval have the same right endpoint $t_2 = a - C_r$, we have

$$x^*(a) = (U - d)C_r, \quad \text{for } x(0) \in [-(U - d)(a - C_r), d(a - C_r)]. \quad (4.98)$$

If the initial production surplus level $x(0)$ is less than $-(U - d)(a - C_r)$, the machine produces at its maximum rate during its time of operation $[0, a]$.

For an initial production surplus level $x(0)$ satisfying $x(0) \in [d(a - C_r), d(a + \frac{b}{2})]$, we have a switching point at time $t_0 = a - S_1(x(0))$.

For an initial production surplus level $x(0)$ satisfying $x(0) \geq d(a + \frac{b}{2})$, the machine is shut down during the entire interval $[0, a]$.

We now examine the trajectories which have a switching point more closely. In order to illustrate the control structure, we define a switching curve $x_S(t_0)$ at which an optimal control switches from a production rate of 0 to U . Before the switching time t_0 , the machine

has been shut down and the production surplus decreases with rate d . Therefore, at the switching point, we have

$$x_S(t_0) = x(0) - dt_0, \quad \text{for } t_0 \in [a - C_r, a]. \quad (4.99)$$

From Equation (4.86), we also have

$$t_0 = a - S_1(x(0)). \quad (4.100)$$

Combining Equations (4.99) and (4.100) leads to

$$S_1(x_S(t_0) + dt_0) = a - t_0, \quad \text{for } t_0 \in [a - C_r, a]. \quad (4.101)$$

From Equation (4.77), we obtain

$$\begin{aligned} & S_1(x_S(t_0) + dt_0) \\ = & -\frac{x_S(t_0) + dt_0 - da + Ub}{U + d} + \sqrt{\left(\frac{x_S(t_0) + dt_0 - da + Ub}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\left(a + \frac{b}{2}\right) - x_S(t_0) - dt_0\right)} \\ = & a - t_0. \end{aligned}$$

By rearranging the terms and taking the square, we conclude

$$\begin{aligned} & \left(\frac{x_S(t_0) + dt_0 - da + Ub}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\left(a + \frac{b}{2}\right) - x_S(t_0) - dt_0\right) \\ = & \left(\frac{x_S(t_0) + dt_0 - da + Ub}{U + d} + a - t_0\right)^2 \\ = & \left(\frac{x_S(t_0) + dt_0 - da + Ub}{U + d}\right)^2 + 2\frac{x_S(t_0) + dt_0 - da + Ub}{U + d}(a - t_0) + (a - t_0)^2. \end{aligned}$$

The term $\left(\frac{x_S(t_0) + dt_0 - da + Ub}{U + d}\right)^2$ cancels out on both sides. Rearranging the remaining terms results in

$$x_S(t_0) = \frac{Ub^2}{2(a + b - t_0)} - \frac{U - d}{2}(a + b - t_0), \quad \text{for } t_0 \in [a - C_r, a]. \quad (4.102)$$

Starting from an initial surplus $x(0) \in [d(a - C_r), d(a + \frac{1}{2}b)]$, the production surplus $x^*(t)$ decreases until the trajectory reaches the switching curve, i.e.,

$$x^*(t_0) = x_S(t_0), \quad \text{for } t_0 \in [a - C_r, a].$$

The production rate is then switched to the maximum level and $x^*(t)$ increases until the end of the uptime interval.

At the endpoints of the interval $[a - C_r, a]$, the switching curve satisfies

$$\begin{aligned}
 x_S(a - C_r) &= \frac{Ub^2}{2(a + b - (a - C_r))} - \frac{U - d}{2}(a + b - (a - C_r)) \\
 &= \frac{1}{2(b + C_r)}(Ub^2 - (U - d)(b + C_r)^2) \\
 &= \frac{1}{2(b + C_r)}(Ub^2 - (U - d)\frac{U}{U - d}b^2) \\
 &= 0,
 \end{aligned}$$

and

$$\begin{aligned}
 x_S(a) &= \frac{Ub^2}{2(a + b - a)} - \frac{U - d}{2}(a + b - a) \\
 &= \frac{db}{2}.
 \end{aligned}$$

Notice also that the switching curve $x_S(t)$ for $t \in (a - C_r, a]$ lies above the trajectory $x_0^*(t)$, since

$$x_S(a - C_r) = x_0^*(a - C_r) = 0, \quad (4.103)$$

and

$$\begin{aligned}
 \frac{d}{dt}x_S(t) &= \frac{Ub^2}{2(a + b - t)^2} + \frac{U - d}{2} \\
 &> \frac{Ub^2}{2(a + b - (a - C_r))^2} + \frac{U - d}{2} \\
 &= \frac{Ub^2}{2(b + C_r)^2} + \frac{U - d}{2} \\
 &= \frac{U - d}{2} + \frac{U - d}{2} \\
 &= (U - d) \\
 &= \frac{d}{dt}x_0^*(t), \quad \text{for } t \in (a - C_r, a). \quad (4.104)
 \end{aligned}$$

This can also be seen in Figure 4.14. Thus, the optimal trajectory for an initial production surplus level $x(0) \in (d(a - C_r), d(a + \frac{b}{2}))$ does not touch the trajectory $x_0^*(t)$.

Since the optimal production policy $u_1^*(t)$ is determined by the initial production surplus level $x(0)$, we can therefore express the optimal performance measure in dependence of $x(0)$ as well. The optimal performance measure $J_1^*(x(0))$ satisfies

$$J_1^*(x(0)) = \left\{ \begin{array}{l} \int_0^a \{x(0) + (U-d)t\}^2 dt + \int_a^{a+b} \{x(0) + Ua - dt\}^2 dt, \\ \qquad \qquad \qquad \text{for } x(0) \leq -(U-d)(a-C_r), \\ \\ \dots\dots\dots \\ \int_0^{-\frac{x(0)}{U-d}} \{x(0) + (U-d)t\}^2 dt + \int_0^{C_r} \{(U-d)t\}^2 dt \\ \qquad \qquad \qquad + \int_0^b \{(U-d)C_r - dt\}^2 dt, \\ \qquad \qquad \qquad \text{for } x(0) \in [-(U-d)(a-C_r), 0], \\ \\ \dots\dots\dots \\ \int_0^{C_r} \{(U-d)t\}^2 dt + \int_0^b \{(U-d)C_r - dt\}^2 dt, \\ \qquad \qquad \qquad \text{for } x(0) = 0, \\ \\ \dots\dots\dots \\ \int_0^{\frac{x(0)}{d}} \{x(0) - dt\}^2 dt + \int_0^{C_r} \{(U-d)t\}^2 dt + \int_0^b \{(U-d)C_r - dt\}^2 dt, \\ \qquad \qquad \qquad \text{for } x(0) \in [0, d(a-C_r)], \\ \\ \dots\dots\dots \\ \int_0^{a-S_1(x(0))} \{x(0) - dt\}^2 dt \\ \qquad \qquad \qquad + \int_0^{S_1(x(0))} \{x(0) - d(a-S_1(x(0))) + (U-d)t\}^2 dt \\ \qquad \qquad \qquad + \int_a^{a+b} \{x(0) + US_1(x(0)) - dt\}^2 dt, \\ \qquad \qquad \qquad \text{for } x(0) \in [d(a-C_r), d(a+\frac{b}{2})], \\ \\ \dots\dots\dots \\ \int_0^{a+b} \{x(0) - dt\}^2 dt, \\ \qquad \qquad \qquad \text{for } x(0) \geq d(a+\frac{b}{2}). \end{array} \right. \quad (4.105)$$

The first derivative of the function $J_1^*(x(0))$ with respect to $x(0)$ according to Appendix C.4 is

$$\frac{d}{dx(0)} J_1^*(x(0)) = \begin{cases} 2x(0)(a+b) + (U-d)(a^2 + 2ab + (b^2 - b^2)) - db^2, & \text{for } x(0) < -(U-d)(a-C_r), \\ \dots\dots\dots \\ -\frac{x^2(0)}{U-d}, & \text{for } x(0) \in (-(U-d)(a-C_r), 0), \\ \dots\dots\dots \\ \frac{x^2(0)}{d}, & \text{for } x(0) \in (0, d(a-C_r)), \\ \dots\dots\dots \\ 2x(0)(a+b) - d(a+b)^2 + US^2(x(0)) + 2US_1(x(0))b & \text{for } x(0) \in (d(a-C_r), d(a+\frac{b}{2})), \\ \dots\dots\dots \\ 2x(0)(a+b) - d(a+b)^2, & \text{for } x(0) > d(a+\frac{b}{2}). \end{cases} \quad (4.106)$$

In Appendix C.7, we verify the following property:

$$\begin{aligned} \frac{d}{dx(0)} J_1^*(x_0) &< 0, & \text{for } x(0) < 0, \\ \frac{d}{dx(0)} J_1^*(x_0) &= 0, & \text{for } x(0) = 0, \\ \frac{d}{dx(0)} J_1^*(x_0) &> 0, & \text{for } x(0) > 0. \end{aligned} \quad (4.107)$$

From Equation (4.107), we conclude

$$\begin{aligned} J_1^*(x_1(0)) &< J_1^*(x_2(0)) & \text{for } 0 \leq x_1(0) < x_2(0), \\ J_1^*(x_1(0)) &< J_1^*(x_2(0)) & \text{for } x_2(0) < x_1(0) \leq 0. \end{aligned} \quad (4.108)$$

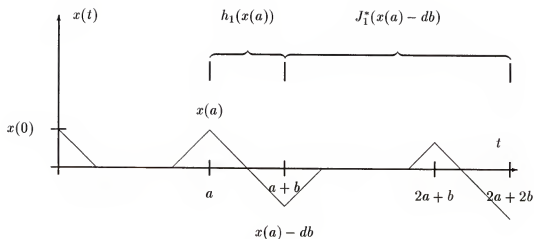


Figure 4.15. A trajectory for the two-cycle problem

at $t = a$ by defining

$$h_2(x(a)) = h_1(x(a)) + J_1^*(x(a) - db), \quad (4.110)$$

where $h_1(x(a))$ is the terminal cost for Problem P_1 with production surplus $x(a)$ and $J_1^*(x(a) - db)$ is the optimal cost for a one-cycle problem starting with a production surplus of $x(a) - db$.

The motivation for the definition of $h_2(x(a))$ is illustrated in Figure 4.15. The term $h_1(x(a))$ reflects the cost during the interval $[a, a + b]$ and the term $J_1^*(x(a) - db)$ reflects the cost during the second cycle, i.e., the interval $[a + b, 2a + 2b]$.

Notice that $h_2(x(a))$ satisfies

$$\frac{d}{dx(a)} h_2(x(a)) = 2x(a)b - db^2 + \frac{d}{dx(a)} J_1^*(x(a) - db). \quad (4.111)$$

The control problem for the two-cycle case can be restated as follows.

Problem P_2 : Determine an admissible control $u^*(t)$, i.e., $u^*(t) \in [0, U]$, for $t \in [0, a]$, that causes the system

$$\frac{dx(t)}{dt} = u(t) - d$$

to follow a trajectory $x^*(t)$ which minimizes the performance measure

$$J_2(u) = \int_0^a x^2(t)dt + h_2(x(a)).$$

The boundary conditions are $x(0) = x_0$ and $x(a)$ free.

For Problem P_2 , the Hamiltonian (see Pontryagin [86], p. 18 and Kirk [67], p. 188) is given by

$$\mathcal{H}(x(t), u(t), p(t)) = x^2(t) + p(t)[u(t) - d], \quad \text{for } t \in [0, a],$$

where the function $p(t)$ is referred to as costate function.

According to Pontryagin's Minimum Principle, necessary conditions for an optimal control law $u^*(t)$ for Problem P_2 (see Pontryagin [86], p. 69 and Kirk [67], p. 233) are

$$\frac{dx^*(t)}{dt} = u^*(t) - d; \quad (4.112)$$

$$\frac{dp^*(t)}{dt} = -2x^*(t); \quad (4.113)$$

$$\begin{aligned} x^{*2}(t) + p^*(t)[u^*(t) - d] &= \min_{u(t) \in [0, U]} \{x^{*2}(t) + p^*(t)[u(t) - d]\} \\ &= x^{*2}(t) - p^*(t)d + \min_{u(t) \in [0, U]} \{p^*(t)u(t)\}. \end{aligned} \quad (4.114)$$

The boundary condition for Problem P_2 satisfies

$$2x^*(a)b - db^2 - p^*(a) + \frac{d}{dx(a)}J_1^*(x(a) - db) = 0. \quad (4.115)$$

Since the final time is fixed and the Hamiltonian does not explicitly depend on time, we have

$$x^{*2}(t) + p^*(t)[u^*(t) - d] = c_1, \quad \text{for } t \in [0, a]. \quad (4.116)$$

From Condition (4.114), the optimal control satisfies

$$u^*(t) = \begin{cases} 0, & \text{if } p^*(t) > 0, \\ U, & \text{if } p^*(t) < 0, \\ \beta, & \text{if } p^*(t) = 0. \end{cases} \quad (4.117)$$

where $\beta \in [0, U]$ is to be determined.

According to Equation (4.117), the optimal control law $u^*(t)$ is thus determined whenever we have $p^*(t) \neq 0$. In the case that roots exist for the function $p^*(t)$ within the interval $[0, a]$, in the same way as in Sections 4.3.1 and 4.3.2 for Problem P_1 , we further distinguish between *singular intervals*, *switching times*, and *switching points*. For Problem P_2 , virtually the same analysis as for Problem P_1 leads to the same results with respect to singular intervals and switching points.

4.4.1 Singular Intervals for Problem P_2

For Problem P_1 , we concluded that if there is a singular interval within $[0, a]$, there is neither a second singular interval nor a switching point within $[0, a]$. This result is based on Equation (4.16). It is independent of the terminal cost. The corresponding Equation for Problem P_2 is Equation (4.116). We conclude that for Problem P_2 , if there is a singular interval within $[0, a]$, there is neither a second singular interval nor a switching point within $[0, a]$.

Next, we demonstrate that Property (4.26) for Problem P_1 can be verified for Problem P_2 as well.

If two admissible controls $u'(t)$ and $u''(t)$ lead to trajectories $x'(t)$ and $x''(t)$

such that (a) $|x'(t)| \leq |x''(t)|$, for $t \in [0, a]$, and

(b) either $0 \geq x'(a) > x''(a)$,

or $db \leq x'(a) < x''(a)$,

then, we have $J_2(u') < J_2(u'')$. (4.118)

To see the validity of this property, notice the following:

1. We have $\int_0^a (x'(t))^2 dt \leq \int_0^a (x''(t))^2 dt$, since $|x'(t)| \leq |x''(t)|$ for $t \in [0, a]$.
2. We have $\int_0^b (x'(a) - dt)^2 dt < \int_0^b (x''(a) - dt)^2 dt$, for $0 \geq x'(a) > x''(a)$ or $x''(a) > x'(a) \geq db$.

3. By Equation (4.108), for $0 \geq x'(a) > x''(a)$ (i.e., $0 > -db \geq x'(a) - db > x''(a) - db$), or $x''(a) > x'(a) \geq da$ (i.e., $0 \leq x'(a) - db < x''(a) - db$), we have $J_1^*(x'(a) - db) < J_1^*(x''(a) - db)$.

Therefore, we can conclude

$$\begin{aligned} J_2(u') &= \int_0^a x'^2(t)dt + \int_0^b (x'(a) - dt)^2dt + J_1^*(x'(a) - db) \\ &\leq \int_0^a x''^2(t)dt + \int_0^b (x''(a) - dt)^2dt + J_1^*(x''(a) - db) = J_2(u''). \end{aligned}$$

In the same way that we use Property (4.26) to conclude Equation (4.27) for Problem P_1 , we use Property (4.118) to conclude for Problem P_2 ,

$$u^*(t) = \begin{cases} \beta_1, & \text{if } t \in [0, t_1), \\ d, & \text{if } t \in [t_1, t_2), \\ U, & \text{if } t \in [t_2, a), \end{cases} \quad (4.119)$$

where $\beta_1 \in \{0, U\}$.

4.4.2 Switching Points for Problem P_2

Based on Property (4.118), we can establish that if the optimal control has a switching point $x^*(t_0)$ at time t_0 , we must have

$$x^*(t_0) \geq 0, \quad (4.120)$$

$$x(0) > 0, \quad (4.121)$$

and

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, t_0), \\ U, & \text{if } t \in [t_0, a). \end{cases} \quad (4.122)$$

Concerning the terms singular interval and switching point, please notice that in our analysis, these terms are only used in reference to the interval $[0, a]$. The remaining time interval can be solved according to the Principle of Optimality and the solution of Problem P_1 .

4.4.3 Initial Control Level for Problem P_2

In Section 4.3.3, we have determined the initial control level for Problem P_1 . We have done so by showing that two of the three potentially optimal control levels 0, d , and U cannot be optimal for certain initial production surplus values. We have distinguished six cases A1 to A6 (see Table 4.2 in Section 4.3.3). We use the same cases for our analysis of the initial control level for Problem P_2 .

In the following, we explain how we arrive at the same result for Problem P_2 as for Problem P_1 .

To show that Cases A1 and A6 are not optimal for Problem P_1 , we have used Property (4.24). Since Property (4.24) is independent of the terminal cost, it holds for Problem P_2 as well. Therefore, for Problem P_2 , Cases A1 and A6 are not optimal either.

For Problem P_1 , Cases A2 and A4 are based on Property (4.26). For Problem P_2 , we have shown the corresponding Property (4.118) above (see Section (4.4.1)). We therefore conclude that Cases A2 and A4 for Problem P_2 are not optimal either.

For Cases A3-1 and A5 in Problem P_1 , we also use Property (4.26). In the same way, the results are transferred for Problem P_2 . The remaining subcase (Case A3-2) needs to be modified for Problem P_2 , since Problems P_1 and P_2 have different terminal cost. The calculations for Case A3-2 for Problem P_2 are shown in Appendix C.9. The result is the same as in Section 4.3.3 for Problem P_1 .

In summary, for Problem P_2 , the initial control level satisfies

$$u^*(0) = \begin{cases} 0, & \text{if } x(0) > 0, \\ d, & \text{if } x(0) = 0, \\ U, & \text{if } x(0) < 0. \end{cases} \quad (4.123)$$

4.4.4 Optimal Solution for Problem P_2

In this section, we derive the optimal control in dependence of the initial production surplus $x(0)$ based on the observations in the previous sections.

Due to the different terminal cost, we partition the range of $x(0)$ for Problem P_2 in a slightly different fashion:

- C1: $x(0) = 0$,
- C2: $x(0) \in (-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), 0)$,
- C3: $x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d})$,
- C3-1: $x(0) \in [-(U-d)a, -(U-d)(a - \frac{1}{2} \frac{db}{U-d})]$,
- C3-2: $x(0) < -(U-d)a$,
- C4: $x(0) \geq d(a - \frac{1}{2} \frac{db}{U-d})$,
- C4-1: $x(0) \geq d(a+b)$,
- C4-2: $x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), d(a+b)]$,
- C4-2-1: $x(0) \in (da + (U - \sqrt{U(U-d)})b, d(a+b))$,
- C4-2-2: $x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b]$,
- C5: $x(0) \in (0, d(a - \frac{1}{2} \frac{db}{U-d}))$.

As we have seen for Problem P_1 , the three possible structures for the optimal control are a Type-1 Structure, a Type-2 Structure, and a Type-3 Structure. For Problem P_2 and each $x(0)$ value, the optimal control has one of these structures as well.

Case C1: $x(0) = 0$

According to Equation (4.123), we have $u^*(0) = d$. Due to Properties (4.34) and (4.36), this excludes a Type-2 Structure and a Type-3 Structure from consideration. Consequently, the optimal trajectory has a Type-1 Structure which has a singular interval $[t_1, t_2]$. Since $u^*(0) = d$, the singular interval has the left endpoint $t_1 = 0$. For the right endpoint, we introduce parameter $C = a - t_2$, so that we have $[t_1, t_2] = [0, a - C]$. From Equation (4.119), we obtain for the production control

$$u(t) = \begin{cases} d, & \text{if } t \in [0, a - C], \\ U, & \text{if } t \in [a - C, a], \end{cases} \quad (4.124)$$

and for the trajectory

$$x(t) = \begin{cases} 0, & \text{if } t \in [0, a - C), \\ (U - d)(t - (a - C)), & \text{if } t \in [a - C, a), \\ (U - d)C - d(t - a), & \text{if } t \in [a, a + b), \end{cases}$$

The performance measure, expressed in terms of C , is

$$\begin{aligned} J_2(C) &= \int_0^a x^2(t)dt + h_2(x(a)) \\ &= \int_0^C ((U - d)t)^2 dt + \int_0^b ((U - d)C - dt)^2 dt + J_1^*((U - d)C - db). \end{aligned} \quad (4.125)$$

We seek to determine the value of $C \in [0, a)$ which solves

$$\min\{J_2(C) : C \in [0, a)\}. \quad (4.126)$$

We differentiate $J_2(C)$ by using Leibniz' rule and obtain

$$\begin{aligned} \frac{d}{dC} J_2(C) &= ((U - d)C)^2 + \int_0^b 2(U - d)((U - d)C - dt)dt + \frac{d}{dC} J_1^*((U - d)C - db) \\ &= (U - d)^2 \{C^2 + 2bC - \frac{d}{U - d} b^2\} + \frac{d}{dC} J_1^*((U - d)C - db). \end{aligned} \quad (4.127)$$

In order to solve (4.126), we first examine the roots of the derivative $\frac{d}{dC} J_2(C)$ within the range $[0, a)$. We consider each term in Equation (4.127) separately.

The first term in Equation (4.127) is identical to $\frac{d}{dC} J_1(C)$ for the one-cycle problem with $x(0) = 0$. According to Equations (4.41) and (4.60), we have

$$\begin{aligned} C^2 + 2bC - \frac{d}{U - d} b^2 &< 0, & \text{for } C \in [0, C_r), \\ C^2 + 2bC - \frac{d}{U - d} b^2 &> 0, & \text{for } C \in (C_r, a), \end{aligned} \quad (4.128)$$

where $C_r = b(\sqrt{\frac{U}{U - d}} - 1)$ is the same constant which optimizes (4.40).

For the second term in Equation (4.127), we use Property (4.107) and conclude

$$\begin{aligned} \frac{d}{dC} J_1^*((U - d)C - db) &< 0, & \text{for } C < \frac{db}{U - d}, \\ \frac{d}{dC} J_1^*((U - d)C - db) &> 0, & \text{for } C > \frac{db}{U - d}. \end{aligned} \quad (4.129)$$

Since

$$C_r = b(\sqrt{\frac{U}{U-d}} - \sqrt{\frac{U-d}{U-d}}) = b \frac{(\sqrt{U})^2 - (\sqrt{U-d})^2}{\sqrt{U-d}(\sqrt{U} + \sqrt{U-d})} < \frac{1}{2} \frac{db}{U-d}, \quad (4.130)$$

and $db \leq (U-d)a$, we have

$$0 < C_r < \frac{db}{U-d} \leq a. \quad (4.131)$$

Therefore, with Equation (4.131), we can conclude from Equations (4.127), (4.128) and (4.129)

$$\begin{aligned} \frac{d}{dC} J_2(C) &< 0, & \text{for } C \in [0, C_r], \\ \frac{d}{dC} J_2(C) &> 0, & \text{for } C \in (\frac{db}{U-d}, a). \end{aligned} \quad (4.132)$$

This implies

$$\begin{aligned} \min\{J_2(C) : C \in [0, C_r]\} &= J_2(C_r), \\ \min\{J_2(C) : C \in [\frac{db}{U-d}, a]\} &= J_2(\frac{db}{U-d}). \end{aligned}$$

Consequently, the solution for (4.126) must be within the interval $[C_r, \frac{db}{U-d}]$. Since we have

$$(U-d)C_r - db \geq (U-d)C_r - (U-d)a = -(U-d)(a - C_r), \quad (4.133)$$

$$C \in [C_r, \frac{db}{U-d}] \quad \text{implies} \quad (U-d)C - db \in [(U-d)C_r - db, 0] \subseteq [-(U-d)(a - C_r), 0]. \quad (4.134)$$

According to Equation (4.106), we have

$$\frac{d}{dx(0)} J_1^*(x(0)) = -\frac{x^2(0)}{U-d}, \quad \text{for } x(0) \in (-(U-d)(a - C_r), 0). \quad (4.135)$$

Applying the chain rule, we therefore obtain for $C \in [C_r, \frac{db}{U-d}]$,

$$\frac{d}{dC} J_1^*((U-d)C - db) = (U-d) \left\{ -\frac{((U-d)C - db)^2}{U-d} \right\} = -(U-d)^2 \left(C - \frac{db}{U-d} \right)^2. \quad (4.136)$$

Substituting Equation (4.136) into Equation (4.127), we obtain

$$\begin{aligned} \frac{d}{dC} J_2(C) &= (U-d)^2 \left\{ C^2 + 2bC - \frac{d}{U-d} b^2 - \left(C - \frac{db}{U-d} \right)^2 \right\} \\ &= (U-d)^2 \left\{ 2b \left(1 + \frac{d}{U-d} \right) C - b \left(1 + \frac{d}{U-d} \right) \frac{d}{U-d} b \right\} \\ &= 2U(U-d)b \left\{ C - \frac{1}{2} \frac{db}{U-d} \right\}, \quad \text{for } C \in [C_r, \frac{db}{U-d}], \end{aligned} \quad (4.137)$$

and

$$\frac{d^2}{dC^2} J_2(C) = 2U(U-d)b, \quad \text{for } C \in [C_r, \frac{db}{U-d}].$$

Therefore, we have a unique minimum at $C = \frac{1}{2} \frac{db}{U-d}$, since

$$\frac{d}{dC} J_2(C) = 0 \quad \text{if and only if} \quad C = \frac{1}{2} \frac{db}{U-d},$$

and

$$\frac{d^2}{dC^2} J_2\left(\frac{1}{2} \frac{db}{U-d}\right) = 2U(U-d)b > 0.$$

Thus, $C = \frac{1}{2} \frac{db}{U-d}$ solves (4.126). This means that the optimal control has the singular interval $[0, a - \frac{1}{2} \frac{db}{U-d}]$, produces at maximum during $[a - \frac{1}{2} \frac{db}{U-d}, a]$, and undergoes maintenance during $[a, a+b]$. At the end of the first cycle, the production surplus level is $x^*(a+b) = -\frac{1}{2}db$.

According to the Principle of Optimality, the optimal trajectory during $[a+b, 2(a+b)]$ for Problem P_2 can be obtained using the optimal control for Problem P_1 emanating from $x(a+b)$. Therefore, during $[a+b, 2a+2b]$, the production is controlled according to Equation (4.97) for an initial surplus of $-\frac{1}{2}db$. Thus, it reaches the zero production surplus level again at $t = a+b + \frac{1}{2} \frac{db}{U-d}$. From $db \leq (U-d)a$ and Equation (4.130), we conclude

$$\frac{1}{2} \frac{db}{U-d} \leq a - \frac{1}{2} \frac{db}{U-d} < a - C_r. \quad (4.138)$$

Thus, the second cycle has the singular interval $[a+b + \frac{1}{2} \frac{db}{U-d}, 2a+b - C_r]$. Within $[2a+b - C_r, 2a+b]$, we produce at maximum capacity, and within $[2a+b, 2a+2b]$, we have

another maintenance interval. In summary, we have for the control law

$$u^*(t) = \begin{cases} d, & \text{if } t \in [0, a - \frac{1}{2} \frac{db}{U-d}), \\ U, & \text{if } t \in [a - \frac{1}{2} \frac{db}{U-d}, a), \\ 0, & \text{if } t \in [a, a+b), \\ U, & \text{if } t \in [a+b, a+b + \frac{1}{2} db), \\ d, & \text{if } t \in [a+b + \frac{1}{2} db, 2a+b - C_r), \\ U, & \text{if } t \in [2a+b - C_r, 2a+b), \\ 0, & \text{if } t \in [2a+b, 2a+2b). \end{cases} \quad (4.139)$$

The optimal trajectory $x^*(t)$ satisfies

$$x^*(t) = \begin{cases} 0, & \text{if } t \in [0, a - \frac{1}{2} \frac{db}{U-d}], \\ (U-d)(t - (a - \frac{1}{2} \frac{db}{U-d})), & \text{if } t \in [a - \frac{1}{2} \frac{db}{U-d}, a], \\ \frac{1}{2} db - d(t - a), & \text{if } t \in [a, a+b], \\ (U-d)(t - (a+b)) - \frac{1}{2} db, & \text{if } t \in [a+b, a+b + \frac{1}{2} \frac{db}{U-d}], \\ 0, & \text{if } t \in [a+b + \frac{1}{2} \frac{db}{U-d}, 2a+b - C_r], \\ (U-d)(t - (2a+b - C_r)), & \text{if } t \in [2a+b - C_r, 2a+b], \\ (U-d)C_r - d(t - (2a+b)), & \text{if } t \in [2a+b, 2a+2b]. \end{cases} \quad (4.140)$$

We have included the second cycle in the expressions for $u^*(t)$ and $x^*(t)$ for illustrative purposes. For the remaining cases C2, C3, C4, and C5, we simply give the optimal production control and the optimal trajectory for the first cycle, since the second cycle for $x(a+b) = x'$ of the two-cycle problem is controlled in the same way as the one-cycle problem for $x(0) = x'$.

We now analyze the costate function. For the costate function $p^*(t)$, we have $p^*(t) = 0$ for $t \in [0, a - \frac{1}{2} \frac{db}{U-d}]$. Using Equation (4.113), we have for $t \in [a - \frac{1}{2} \frac{db}{U-d}, a]$,

$$\begin{aligned} p^*(t) &= p^*(t) - p^*(a - \frac{1}{2} \frac{db}{U-d}) = -2 \int_{a - \frac{1}{2} \frac{db}{U-d}}^t x(\tau) d\tau \\ &= -2 \int_{a - \frac{1}{2} \frac{db}{U-d}}^t (U-d)(\tau - (a - \frac{1}{2} \frac{db}{U-d})) d\tau \\ &= -2 \int_0^{t - (a - \frac{1}{2} \frac{db}{U-d})} (U-d)\tau d\tau = -(U-d)(t - (a - \frac{1}{2} \frac{db}{U-d}))^2. \end{aligned}$$

In summary, we have

$$p^*(t) = \begin{cases} 0, & \text{if } t \in [0, a - \frac{1}{2} \frac{db}{U-d}], \\ -(U-d)(t - (a - \frac{1}{2} \frac{db}{U-d}))^2, & \text{if } t \in [a - \frac{1}{2} \frac{db}{U-d}, a]. \end{cases} \quad (4.141)$$

We verify the boundary condition (see Equation (4.115)):

$$\begin{aligned} &2x^*(a)b - db^2 - p^*(a) + \frac{d}{dx(a)} J_1^*(x^*(a) - db) \\ &= 2(\frac{1}{2}db)b - db^2 - (-(U-d)(\frac{1}{2} \frac{db}{U-d})^2) - \frac{(-\frac{1}{2}db)^2}{U-d} \\ &= 0. \end{aligned} \quad (4.142)$$

This concludes the analysis for Case C1 where $x(0) = 0$. As for the one-cycle problem, before we proceed with the cases where the initial production surplus is negative or positive, we state an optimality condition.

Optimality Condition

If $x_1^*(t)$ and $x_2^*(t)$ are two optimal trajectories for Problem P_2 with initial production surpluses $x_1^*(0)$ and $x_2^*(0)$, respectively, then the trajectories will not cross each other, i.e.,

$$\text{if } x_1^*(t') < x_2^*(t') \text{ for some } t' \in [0, a], \quad \text{then } x_1^*(t) \leq x_2^*(t) \text{ for all } t \in [0, a], \quad (4.143)$$

and in particular for $t' = 0$,

$$\text{if } x_1^*(0) < x_2^*(0), \quad \text{then} \quad x_1^*(t) \leq x_2^*(t), \quad \text{for } t \in [0, a]. \quad (4.144)$$

Case C2: $x(0) \in (-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0)$

First, since $x(0) < 0$, the initial control level satisfies $u^*(0) = U$ according to Equation (4.123). We now examine the control structure. Since $u^*(0) = U$, we exclude a Type-2 Structure according to Equation (4.34). The optimal trajectory for $x(0) \in (-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0)$ meets the trajectory $x_0^*(t)$ at time $t = \frac{-x(0)}{U-d}$. According to the Principle of Optimality, these trajectories cannot cross each other and we conclude that the optimal control has the singular interval $[\frac{-x(0)}{U-d}, a - \frac{1}{2}\frac{db}{U-d}]$. Therefore, we have for the optimal control $u^*(t)$

$$u^*(t) = \begin{cases} U, & \text{if } t \in [0, \frac{-x(0)}{U-d}), \\ d, & \text{if } t \in [\frac{-x(0)}{U-d}, a - \frac{1}{2}\frac{db}{U-d}), \\ U, & \text{if } t \in [a - \frac{1}{2}\frac{db}{U-d}, a]. \end{cases} \quad (4.145)$$

The optimal trajectory $x^*(t)$ satisfies

$$x^*(t) = \begin{cases} (U-d)(t - \frac{-x(0)}{U-d}), & \text{if } t \in [0, \frac{-x(0)}{U-d}], \\ 0, & \text{if } t \in [\frac{-x(0)}{U-d}, a - \frac{1}{2}\frac{db}{U-d}], \\ (U-d)(t - (a - \frac{1}{2}\frac{db}{U-d})), & \text{if } t \in [a - \frac{1}{2}\frac{db}{U-d}, a], \\ \frac{1}{2}db - d(t - a), & \text{if } t \in [a, a + b]. \end{cases} \quad (4.146)$$

For the costate function, we obtain for $t \in [0, \frac{-x(0)}{U-d}]$ according to Equation (4.113),

$$\begin{aligned} p^*(t) &= \int_t^{\frac{-x(0)}{U-d}} 2x^*(\tau) d\tau + p^*\left(\frac{-x(0)}{U-d}\right) \\ &= \int_t^{\frac{-x(0)}{U-d}} 2(U-d)(\tau - \frac{-x(0)}{U-d}) d\tau \\ &= -(U-d)\left(t - \frac{-x(0)}{U-d}\right)^2. \end{aligned} \quad (4.147)$$

Since the trajectory $x^*(t)$ for $t \in [\frac{-x(0)}{U-d}, a + b]$ is identical to the trajectory in Equation (4.140), the remaining interval $[\frac{-x(0)}{U-d}, a]$ yields the same expression for $p^*(t)$ as in Equation

(4.141). In summary, we have

$$p^*(t) = \begin{cases} -(U-d)(t - \frac{-x(0)}{U-d})^2, & \text{if } t \in [0, \frac{-x(0)}{U-d}], \\ 0, & \text{if } t \in [\frac{-x(0)}{U-d}, a - \frac{1}{2}\frac{db}{U-d}], \\ -(U-d)(t - (a - \frac{1}{2}\frac{db}{U-d}))^2, & \text{if } t \in [a - \frac{1}{2}\frac{db}{U-d}, a]. \end{cases} \quad (4.148)$$

Since the expressions for $x^*(a)$ and $p^*(a)$ are the same as in Case C1, the boundary condition (see Equation (4.115)) is verified in the same way as in Equation (4.142). This concludes the considerations for the Case C2 with $x(0) \in (-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0)$.

Case C3: $x(0) < -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$

First, since $x(0) < 0$, the initial control level satisfies $u^*(0) = U$ according to Equation (4.123). Next, we examine the control structure. According to Property 4.34, we can exclude a Type-2 Structure (a production control with a switching point), since it requires $u^*(0) = 0$. Therefore, the optimal control has either a Type-1 Structure or a Type-3 Structure. That is, the optimal control either produces at maximum capacity or it has a singular interval. By distinguishing between the cases $x(0) \in [-(U-d)a, -(U-d)(a - \frac{1}{2}\frac{db}{U-d})]$ (Case C3-1) and $x(0) < -(U-d)a$ (Case C3-2), we show that the optimal control has a Type-3 Structure.

Case C3-1: $x(0) \in [-(U-d)a, -(U-d)(a - \frac{1}{2}\frac{db}{U-d})]$.

If we have a Type-3 Structure, we produce at *maximum capacity*,

$$u(t) = U, \quad \text{for } t \in [0, a]. \quad (4.149)$$

If we have a Type-1 Structure, we have *singular interval* $[t_1, t_2]$ and the production control satisfies Equation (4.119). Since we have $u^*(0) = U$, we conclude $\beta_1 = U$ in Equation (4.119). The left endpoint of the singular interval satisfies $t_1 = \frac{-x(0)}{U-d}$. Let $t_2 = a - C$ where $C \in [0, a - \frac{-x(0)}{U-d}]$. With $[t_1, t_2] = [\frac{-x(0)}{U-d}, a - C]$, the production control thus satisfies

$$u(t) = \begin{cases} U, & \text{if } t \in [0, \frac{-x(0)}{U-d}), \\ d, & \text{if } t \in [\frac{-x(0)}{U-d}, a - C], \\ U, & \text{if } t \in [a - C, a]. \end{cases} \quad (4.150)$$

Notice that we can represent the production control in Equation (4.149), by letting $C = a - \frac{-x(0)}{U-d}$ in Equation (4.154). Thus, if we extend the range of C from $[0, a - \frac{-x(0)}{U-d})$ to $[0, a - \frac{-x(0)}{U-d}]$ in Equation (4.150), we have represented both control structures which need to be considered.

We can express the performance measure as a function of the parameter C ,

$$J_2(C) = \int_0^{\frac{-x(0)}{U-d}} ((U-d)t)^2 dt + \int_0^C ((U-d)t)^2 dt + \int_0^b ((U-d)C - dt)^2 dt + J_1^*((U-d)C - db). \quad (4.151)$$

We seek to determine the value of C which solves

$$\min\{J_2(C) : C \in [0, a - \frac{-x(0)}{U-d}]\}. \quad (4.152)$$

Notice that the cost function $J_2(C)$ in Equation (4.151) differs from the cost function in Equation (4.125) only by the term $\int_0^{\frac{-x(0)}{U-d}} ((U-d)t)^2 dt$. This term is independent of C . Therefore, the derivatives of the cost functions in Equations (4.151) and (4.125) with respect to C are identical (see Equation (4.127)):

$$\frac{d}{dC} J_2(C) = (U-d)^2 \{C^2 + 2bC - \frac{d}{U-d} b^2\} + \frac{d}{dC} J_1^*((U-d)C - db).$$

With the same reasoning as in Case C1, we obtain

$$\frac{d}{dC} J_2(C) = \begin{cases} < 0 & \text{for } C \in [0, C_r), \\ 2U(U-d)b(C - \frac{1}{2}\frac{db}{U-d}) & \text{for } C \in [C_r, \frac{db}{U-d}], \\ > 0 & \text{for } C \in (\frac{db}{U-d}, a]. \end{cases} \quad (4.153)$$

Since we have

$$\begin{aligned} a - \frac{-x(0)}{U-d} &\leq a + \frac{1}{U-d} (-(U-d)(a - \frac{1}{2}\frac{db}{U-d})) \\ &= \frac{1}{2}\frac{db}{U-d}, \quad \text{for } x(0) \leq -(U-d)(a - \frac{1}{2}\frac{db}{U-d}), \end{aligned}$$

we conclude from Equation (4.153)

$$\frac{d}{dC} J_2(C) < 0, \quad \text{for } C \in [0, a - \frac{-x(0)}{U-d}),$$

and thus, $C = a - \frac{x(0)}{U-d}$ solves (4.152). Equation (4.150) with $C = a - \frac{x(0)}{U-d}$ leads to

$$u^*(t) = U, \quad \text{for } t \in [0, a]. \quad (4.154)$$

Therefore, we have shown that the optimal control for $x(0) \in [-(U-d)a, -(U-d)(a - \frac{1}{2}\frac{db}{U-d})]$ is a Type-3 Structure.

Case C3-2: $x(0) < -(U-d)a$.

Recall that we have already excluded a Type-2 Structure. In this case, we can also exclude a Type-1 Structure, since the trajectory $x(t)$ cannot reach the zero production surplus level for $t \in [0, a]$ and $x(0) < -(U-d)a$.

Thus, we have a Type-3 Structure, and since the initial control level satisfies $u^*(0) = U$, we have the same optimal control as in Case C3-1 (Equation (4.154)).

For the optimal trajectory in both Cases C3-1 and C3-2, we have

$$x^*(t) = \begin{cases} x(0) + (U-d)t, & \text{if } t \in [0, a], \\ x(0) + (U-d)a - d(t-a), & \text{if } t \in [a, a+b]. \end{cases} \quad (4.155)$$

For the costate function, we use Equation (4.113), to obtain $p^*(a) - p^*(t) = \int_t^a -2x^*(\tau)d\tau$, or equivalently,

$$\begin{aligned} p^*(t) &= 2 \int_t^a x^*(\tau)d\tau + p^*(a), \\ &= 2 \int_t^a (x(0) + (U-d)\tau)d\tau + p^*(a), \\ &= 2x(0)(a-t) + (U-d)(a^2 - t^2) + p^*(a), \quad \text{for } t \in [0, a]. \end{aligned} \quad (4.156)$$

The boundary condition (Equation (4.115)) results in an expression for $p^*(a)$:

$$p^*(a) = 2x^*(a)b - db^2 + \frac{d}{dx(a)}J_1^*(x^*(a) - db), \quad (4.157)$$

where $\frac{d}{dx(a)}J_1^*(x^*(a) - db)$ is given by Equation (4.106).

Together with Equations (4.156) and (4.157), we obtain

$$p^*(t) = 2(a+b-t)x(0) + (U-d)(a^2 + 2ab - t^2) - db^2 + \frac{d}{dx(a)}J_1^*(x^*(a) - db). \quad (4.158)$$

This completes the analysis of the case $x(0) \leq -(U - d)(a - \frac{1}{2} \frac{db}{U-d})$.

Cases C1, C2, and C3 complete the analysis of $x(0) \leq 0$. In the following, we consider the cases where $x(0) > 0$.

Case C4: $x(0) > d(a - \frac{1}{2} \frac{db}{U-d})$.

First, since $x(0) > 0$, the initial control level satisfies $u^*(0) = 0$ according to Equation (4.123). Next, with similar reasoning as for Case B4 in Section 4.3.4, we can exclude a Type-1 Structure from consideration.

Thus, we have either a Type-2 Structure or a Type-3 Structure: Since we have $u^*(0) = 0$ for $x(0) > 0$ according to Equation (4.123), we either *shut* the machine *down* during the entire interval or we switch the production rate from 0 to U at some time $t_0 \in (0, a)$.

If we *shut* the machine *down*, the control policy is

$$u(t) = 0, \quad \text{for } t \in [0, a). \quad (4.159)$$

If we have a *switching point* policy with a switching time $t_0 \in (0, a)$, we let $S = a - t_0 \in (0, a)$. According to Equation (4.122), we have with switching time $t_0 = a - S$,

$$u(t) = \begin{cases} 0, & \text{if } t \in [0, a - S), \\ U, & \text{if } t \in [a - S, a). \end{cases} \quad (4.160)$$

We now combine the solution structures by extending the range of S in Equation (4.160) from $(0, a)$ to $[0, a)$, since the production policy in Equation (4.159) is represented by the policy in Equation (4.160) with $S = 0$.

In order to obtain the performance measure, we consider the two cycles separately. For the first cycle, we obtain the same expression as in Case B4 (see Equation (4.73)).

Since we have

$$x^*(a + b) = x^*(a) - db = x(0) + US - d(a + b),$$

the cost for the second cycle is given by $J_1^*(x(0) + US - d(a + b))$. Therefore, the cost function $J_2(S)$ satisfies

$$\begin{aligned} J_2(S) &= \int_0^{a-S} (x(0) - dt)^2 dt + \int_0^S (x(0) - d(a - S) + (U - d)t)^2 dt \\ &\quad + \int_0^b (x(0) - da + US - dt)^2 dt + J_1^*(x(0) + US - d(a + b)). \end{aligned} \quad (4.161)$$

We are seeking the value of parameter S which solves

$$\min\{J_2(S) : S \in [0, a]\}. \quad (4.162)$$

For the derivative, we have according to Leiniz' rule

$$\begin{aligned} \frac{d}{dS} J_2(S) &= -1(x(0) - d(a - S))^2 + (x(0) - d(a - S) + (U - d)S)^2 \\ &\quad + \int_0^S 2d(x(0) - d(a - S) + (U - d)t)dt + \int_0^b 2U(x(0) - da + US - dt)dt \\ &\quad + \frac{d}{dS} J_1^*(x(0) + US - d(a + b)). \end{aligned}$$

Evaluating the integrals and rearranging the terms, we obtain

$$\begin{aligned} \frac{d}{dS} J_2(S) &= U(U + d)\{S^2 + 2\frac{x(0) - da + Ub}{U + d}S - \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))\} \\ &\quad + \frac{d}{dS} J_1^*(x(0) + US - d(a + b)). \end{aligned} \quad (4.163)$$

In the following, we distinguish the cases $x(0) \geq d(a + b)$ (Case C4-1) and $x(0) \in [d(a - \frac{1}{2}\frac{db}{U-d}), d(a + b)]$ (Case C4-2).

Case C4-1: $x(0) \geq d(a + b)$.

We have according to Property (4.107)

$$\frac{d}{dS} J_1^*(x(0) + US - d(a + b)) \geq 0, \quad \text{for } x(0) \geq d(a + b) \text{ and } S \in [0, a].$$

According to Equation (4.163), we conclude

$$\frac{d}{dS} J_2(S) > 0, \quad \text{for } x(0) \geq d(a + b) \text{ and } S \in [0, a].$$

Consequently, $S = 0$ solves (4.162). This leads to the control

$$u^*(t) = 0, \quad \text{for } t \in [0, a], \quad (4.164)$$

and the trajectory

$$x^*(t) = x(0) - dt, \quad \text{for } t \in [0, a + b]. \quad (4.165)$$

With $p^*(a) = 2x^*(a)b - db^2 + \frac{d}{dx(a)}J_1^*(x(a) - db)$ from the boundary condition (see Equation (4.115)), Equation (4.113) for $t \in [0, a]$ leads to

$$\begin{aligned} p^*(t) &= 2 \int_t^a x^*(\tau) d\tau + p^*(a) \\ &= 2 \int_t^a (x(0) - d\tau) d\tau + 2(x(0) - da)b - db^2 + \frac{d}{dx(a)}J_1^*(x(a) - db) \\ &= 2x(0)(a - t) - d(a^2 - t^2) + 2(x(0) - da)b - db^2 + \frac{d}{dx(a)}J_1^*(x(a) - db) \\ &= 2x(0)(a + b - t) - d((a + b)^2 - t^2) + \frac{d}{dx(a)}J_1^*(x(0) - d(a + b)). \end{aligned} \quad (4.166)$$

This completes our analysis for $x(0) \geq d(a + b)$.

Case C4-2: $x(0) \in [d(a - \frac{1}{2}\frac{db}{U-d}), d(a + b))$.

In order to evaluate the right-hand-side expression for the derivative in Equation (4.163), we need to evaluate the term $\frac{d}{dS}J_1^*(x(0) + US - d(a + b))$. According to Equation (4.106), we therefore need to identify the possible range of $x(a + b) = x(0) + US - d(a + b)$. From Cases C1 and C4-1, we have

$$x^*(a + b) = -\frac{1}{2}db, \quad \text{for } x(0) = 0, \quad (4.167)$$

$$x^*(a + b) = 0, \quad \text{for } x(0) = d(a + b). \quad (4.168)$$

Together with the Optimality Condition (see Equation (4.144)), we conclude from Equations (4.167) and (4.168)

$$x(a + b) = x(0) + US - d(a + b) \in [-\frac{1}{2}db, 0], \quad \text{for } x(0) \in [0, d(a + b)]. \quad (4.169)$$

Since we have $-\frac{1}{2}db > -(U - d)(a - C_r)$ (see Equation (4.138)), we also have

$$[-\frac{1}{2}db, 0] \subset [-(U - d)(a - C_r), 0]. \quad (4.170)$$

From Equations (4.169) and (4.170), we obtain the following expression according to Equation (4.106),

$$\frac{d}{dS}J_1^*(x(0) + US - d(a + b)) = U(-\frac{(x(0) + US - d(a + b))^2}{U - d})$$

$$= -\frac{U}{U-d}(x(0) + US - d(a+b))^2. \quad (4.171)$$

We can now substitute Equation (4.171) into Equation (4.161) and obtain

$$\begin{aligned} \frac{d}{dS} J_2(S) = & U(U+d)\{S^2 + 2\frac{x(0) - da + Ub}{U+d}S - \frac{2b}{U+d}(d(a + \frac{b}{2}) - x(0))\} \\ & - \frac{U}{U-d}(x(0) + US - d(a+b))^2. \end{aligned} \quad (4.172)$$

According to Appendix C.10, this equation can be rewritten as follows:

$$\begin{aligned} \frac{d}{dS} J_2(S) = & -\frac{Ud^2}{U-d}\{S^2 - 2\frac{U^2b - d(x(0) - da)}{d^2}S \\ & + \frac{x^2(0) - 2(Ub + da)x(0) + d(da^2 + 2Uab + Ub^2)}{d^2}\}. \end{aligned} \quad (4.173)$$

Consequently, the roots of $\frac{d}{dS} J_2(S)$ are

$$\begin{aligned} S'_2 &= \frac{U^2b - d(x(0) - da)}{d^2} - \sqrt{\left(\frac{U^2b - d(x(0) - da)}{d^2}\right)^2 - f(x(0))}, \\ S''_2 &= \frac{U^2b - d(x(0) - da)}{d^2} + \sqrt{\left(\frac{U^2b - d(x(0) - da)}{d^2}\right)^2 - f(x(0))}, \end{aligned}$$

$$\text{where } f(x(0)) = \frac{x^2(0) - 2(Ub + da)x(0) + d(da^2 + 2Uab + Ub^2)}{d^2}.$$

For the second derivative, we have

$$\frac{d^2}{dS^2} J_2(S) = -\frac{Ud^2}{U-d}\{2S - 2\frac{U^2b - d(x(0) - da)}{d^2}\},$$

and therefore,

$$\begin{aligned} \frac{d^2}{dS^2} J_2(S'_2) &> 0, \\ \frac{d^2}{dS^2} J_2(S''_2) &< 0. \end{aligned}$$

Since we are minimizing, we focus on S'_2 and rename it as $S_2(x(0))$:

$$S_2(x(0)) = \frac{U^2b - d(x(0) - da)}{d^2} - \sqrt{\left(\frac{U^2b - d(x(0) - da)}{d^2}\right)^2 - f(x(0))}. \quad (4.174)$$

The unrestricted minimum is therefore at $S = S_2(x(0))$:

$$\min\{J_2(S) : S \text{ unrestricted}\} = J_2(S_2(x(0))). \quad (4.175)$$

Since the range for S in (4.162) is restricted to $[0, a]$, we distinguish the cases $S_2(x(0)) < 0$ and $S_2(x(0)) \geq 0$. According to Equation (C.69) in Appendix C.10, we have that $x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), d(a+b))$ and $S_2(x(0)) < 0$ imply

$$x(0) \in (da + (U - \sqrt{U(U-d)})b, d(a+b)),$$

and $x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), d(a+b))$ and $S_2(x(0)) \geq 0$ imply

$$x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b].$$

Thus, we have Cases C4-2-1 ($x(0) \in (da + (U - \sqrt{U(U-d)})b, d(a+b))$) and C4-2-2 ($x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b]$). For Case C4-2-1, we will obtain a Type-3 Structure, and for Case C4-2-2, we will obtain a Type-2 Structure.

Case C4-2-1: $x(0) \in (da + (U - \sqrt{U(U-d)})b, d(a+b))$.

According to the quadratic function in Equation (4.173), we have

$$\frac{d}{dS} J_2(S) > 0 \quad \text{if and only if} \quad S \in (S_2(x(0)), S_2''). \quad (4.176)$$

Since $S_2(x(0)) < 0$ and the range in (4.162) is $[0, a]$, we conclude from Equation (4.176),

$$\frac{d}{dS} J_2(S) > 0, \quad \text{for } S \in [0, S_2'').$$

Consequently, $S = 0$ solves (4.162). The optimal production control, the optimal trajectory, and the costate function are therefore the same as in Case C4-1.

Case C4-2-2: $x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b]$.

According to Appendices C.11 and C.12, we have

$$S_2(x(0)) \leq a - \frac{1}{2} \frac{db}{U-d} < a, \quad \text{for } x(0) \in (d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b).$$

Due to Equation (4.175), $S = S_2(x(0))$ solves (4.162). Thus, we have a Type-2 Structure with a switching point at $t_0 = a - S_2(x(0))$. The optimal control policy satisfies

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, a - S_2(x(0))], \\ U, & \text{if } t \in [a - S_2(x(0)), a]. \end{cases} \quad (4.177)$$

For the trajectory $x^*(t)$, we have

$$x^*(t) = \begin{cases} x(0) - dt, & \text{if } t \in [0, a - S_2(x(0))], \\ x(0) - dt + U(t - (a - S_2(x(0)))) & \text{if } t \in [a - S_2(x(0)), a], \\ x(0) - dt + US_2(x(0)), & \text{if } t \in [a, a + b]. \end{cases} \quad (4.178)$$

For the function $p^*(t)$, we have a root at the switching point, i.e., $p^*(a - S_2(x(0))) = 0$.

Together with Equation (4.113), we derive

$$p^*(t) = \begin{cases} \int_t^{a-S_2(x(0))} 2x^*(\tau) d\tau, & \text{if } t \in [0, a - S_2(x(0))], \\ \int_{a-S_2(x(0))}^t -2x^*(\tau) d\tau, & \text{if } t \in [a - S_2(x(0)), a]. \end{cases}$$

For $t \in [0, a - S_2(x(0))]$, this yields

$$\begin{aligned} p^*(t) &= \int_t^{a-S_2(x(0))} 2x^*(\tau) d\tau = \int_t^{a-S_2(x(0))} 2(x(0) - d\tau) d\tau \\ &= 2x(0)(a - S_2(x(0)) - t) - d((a - S_2(x(0)))^2 - t^2). \end{aligned}$$

For $t \in [a - S_2(x(0)), a]$, this yields

$$\begin{aligned} &p^*(t) \\ &= \int_{a-S_2(x(0))}^t -2x^*(\tau) d\tau \\ &= -2 \int_{a-S_2(x(0))}^t \{x(0) - d(a - S_2(x(0))) \\ &\quad + (U - d)(\tau - (a - S_2(x(0))))\} d\tau \\ &= -2 \int_{a-S_2(x(0))}^t \{x(0) - U(a - S_2(x(0))) + (U - d)\tau\} d\tau \\ &= -2(x(0) - U(a - S_2(x(0))))(t - (a - S_2(x(0)))) - (U - d)(t^2 - (a - S_2(x(0)))^2). \end{aligned}$$

In summary, we have

$$p^*(t) = \begin{cases} 2x(0)(a - S_2(x(0)) - t) - d((a - S_2(x(0)))^2 - t^2), \\ \quad \text{if } t \in [0, a - S_2(x(0))]; \\ -2(x(0) - U(a - S_2(x(0))))(t - (a - S_2(x(0)))) \\ \quad - (U - d)(t^2 - (a - S_2(x(0)))^2), \\ \quad \text{if } t \in [a - S_2(x(0)), a]. \end{cases} \quad (4.179)$$

Finally, we verify the boundary condition (4.115). With $x^*(a) = x(0) - da + U S_2(x(0))$,

$$\begin{aligned}
 & p^*(a) \\
 = & -2(x(0) - U(a - S_2(x(0))))(a - (a - S_2(x(0)))) - (U - d)(a^2 - (a - S_2(x(0)))^2) \\
 = & -2(x(0) - U(a - S_2(x(0))))S_2(x(0)) - (U - d)(2aS_2(x(0)) - S_2(x(0))^2) \\
 = & -(U + d)(S_2(x(0)))^2 - 2(x(0) - da)S_2(x(0)),
 \end{aligned}$$

and

$$\frac{d}{dx(a)} J_1^*(x(a) - db) = -\frac{(x(a) - db)^2}{U - d} = -\frac{(x(0) - d(a + b) + U S_2(x(0)))^2}{U - d},$$

we have according to Equations (4.173) and (4.174)

$$\begin{aligned}
 & 2x^*(a)b - db^2 - p^*(a) + \frac{d}{dx(0)} J_1^*(x(a) - db) \\
 = & 2\{x(0) - da + U S_2(x(0))\}b - db^2 + (U + d)(S_2(x(0)))^2 + 2(x(0) - da)S_2(x(0)) \\
 & - \frac{(x(0) - d(a + b) + U S_2(x(0)))^2}{U - d} \\
 = & 2x(0)b - 2dab + 2U S_2(x(0))b - db^2 \\
 & + (U + d)(S_2(x(0)))^2 + 2x(0)S_2(x(0)) - 2daS_2(x(0)) \\
 & - \frac{(x(0) - d(a + b))^2 + 2(x(0) - d(a + b))U S_2(x(0)) + U^2(S_2(x(0)))^2}{U - d} \\
 = & \frac{1}{U - d}\{2(U - d)x(0)b - 2(U - d)dab + 2U(U - d)S_2(x(0))b - (U - d)db^2 \\
 & + (U^2 - d^2)(S_2(x(0)))^2 + 2(U - d)x(0)S_2(x(0)) - 2(U - d)daS_2(x(0)) \\
 & - (x(0) - d(a + b))^2 - 2(x(0) - d(a + b))U S_2(x(0)) - U^2(S_2(x(0)))^2\} \\
 = & \frac{1}{U - d}\{2(U - d)x(0)b - 2(U - d)dab + 2U^2 S_2(x(0))b - (U - d)db^2 \\
 & - d^2(S_2(x(0)))^2 - 2dx(0)S_2(x(0)) + 2d^2aS_2(x(0)) \\
 & - (x(0) - d(a + b))^2\} \\
 = & \frac{1}{U - d}\{-d^2(S_2(x(0)))^2 + 2(U^2b - d(x(0) - da))S_2(x(0)) \\
 & - (x(0) - d(a + b))^2 + 2(U - d)bx(0) - (U - d)(2dab + db^2)\} \\
 = & \frac{1}{U - d}\{-d^2(S_2(x(0)))^2 + 2(U^2b - d(x(0) - da))S_2(x(0))
 \end{aligned}$$

$$\begin{aligned}
& -x^2(0) + 2d(a+b)x(0) - d^2(a+b)^2 + 2(U-d)bx(0) - (U-d)(2dab+db^2)\} \\
= & \frac{1}{U-d} \{-d^2(S_2(x(0)))^2 + 2(U^2b - d(x(0)-da))S_2(x(0)) \\
& -x^2(0) + 2(Ub+da)x(0) - d^2(a^2+2ab+b^2) - d(U-d)(2ab+b^2)\} \\
= & \frac{1}{U-d} \{-d^2(S_2(x(0)))^2 + 2(U^2b - d(x(0)-da))S_2(x(0)) \\
& -x^2(0) + 2(Ub+da)x(0) - d^2a^2 - Ud(2ab+b^2)\} \\
= & 0.
\end{aligned} \tag{4.180}$$

Thus, the boundary condition is verified.

This completes the analysis for the case $S_2(x(0)) \geq 0$ which corresponds to the situation $x(0) \in [d(a - \frac{1}{2}\frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b]$.

Case C5: $x(0) \in (0, d(a - \frac{1}{2}\frac{db}{U-d}))$

First, since $x(0) > 0$, the initial control level satisfies $u^*(0) = 0$ according to Equation (4.123). Concerning the optimal control structure, we eliminate a Type-2 and a Type-3 Structure with similar reasoning as for Case B5 in Section 4.3.4. Therefore, we must have a Type-1 Structure which has a singular interval.

The left endpoint satisfies $t_1 = \frac{x(0)}{d}$. At time t_1 , the trajectory for $x(0) \in (0, d(a - \frac{1}{2}\frac{db}{U-d}))$ meets the trajectory $x_0^*(t)$. By the Principle of Optimality, these trajectories coincide from then on. Therefore, the right endpoint is $t_2 = a - \frac{1}{2}\frac{db}{U-d}$.

Thus, the optimal control has the singular interval $[\frac{x(0)}{d}, a - \frac{1}{2}\frac{db}{U-d}]$ and satisfies

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, \frac{x(0)}{d}), \\ d, & \text{if } t \in [\frac{x(0)}{d}, a - \frac{1}{2}\frac{db}{U-d}), \\ U, & \text{if } t \in [a - \frac{1}{2}\frac{db}{U-d}, a). \end{cases} \tag{4.181}$$

The optimal trajectory $x^*(t)$ satisfies

$$x^*(t) = \begin{cases} d(\frac{x(0)}{d} - t), & \text{if } t \in [0, \frac{x(0)}{d}], \\ 0, & \text{if } t \in [\frac{x(0)}{d}, a - \frac{1}{2}\frac{db}{U-d}], \\ (U-d)(t - (a - \frac{1}{2}\frac{db}{U-d})), & \text{if } t \in [a - \frac{1}{2}\frac{db}{U-d}, a], \\ (U-d)\frac{1}{2}\frac{db}{U-d} - d(t-a), & \text{if } t \in [a, a+b]. \end{cases} \quad (4.182)$$

Concerning the function $p^*(t)$, we must have $p^*(\frac{x(0)}{d}) = 0$ and $p^*(\frac{x(0)}{d}) - p^*(t) = -p^*(t) = \int_t^{\frac{x(0)}{d}} -2x^*(\tau)d\tau$. Therefore, we have for $t \in [0, \frac{x(0)}{d}]$,

$$p^*(t) = \int_t^{\frac{x(0)}{d}} 2x^*(\tau)d\tau = \int_t^{\frac{x(0)}{d}} 2d(\frac{x(0)}{d} - \tau)d\tau = -d(\frac{x(0)}{d} - t)^2 \Big|_t^{\frac{x(0)}{d}} = d(\frac{x(0)}{d} - t)^2.$$

The remaining interval $[\frac{x(0)}{d}, a]$ results in the same expression for $p^*(t)$ as in Cases C1 and C2 (see Equations (4.141) and (4.148)), since the trajectory $x^*(t)$ for $t \in [a - \frac{1}{2}\frac{db}{U-d}, a+b]$ is identical to the optimal trajectories in Cases C1 and C2.

In summary, we have

$$p^*(t) = \begin{cases} d(\frac{x(0)}{d} - t)^2, & \text{if } t \in [0, \frac{x(0)}{d}], \\ 0, & \text{if } t \in [\frac{x(0)}{d}, a - \frac{1}{2}\frac{db}{U-d}], \\ -(U-d)(t - (a - \frac{1}{2}\frac{db}{U-d}))^2, & \text{if } t \in [a - \frac{1}{2}\frac{db}{U-d}, a]. \end{cases} \quad (4.183)$$

Since the values $x^*(a)$ and $p^*(a)$ are the same as in Cases C1 and C2, the boundary condition is verified in the same way as in Cases C1 and C2.

This completes our discussion of the case $x(0) \in (0, d(a - \frac{1}{2}\frac{db}{U-d}))$.

Characterization of the optimal solution

Summarizing Equations (4.139), (4.145), (4.154), (4.164), (4.177), and (4.181), we have

$$\begin{aligned}
 u_2^*(t) = & \left\{ \begin{array}{ll} U, & \text{if } x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d}), \text{ and } t \in [0, a); \\ \\ \dots\dots\dots \\ U, & \text{if } x(0) \in (-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), 0), \text{ and } t \in [0, \frac{-x(0)}{U-d}), \\ d, & \text{if } x(0) \in (-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), 0), \text{ and } t \in [\frac{-x(0)}{U-d}, a - \frac{1}{2} \frac{db}{U-d}), \\ U, & \text{if } x(0) \in (-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), 0), \text{ and } t \in [a - \frac{1}{2} \frac{db}{U-d}, a); \\ \\ \dots\dots\dots \\ d, & \text{if } x(0) = 0, \text{ and } t \in [0, a - \frac{1}{2} \frac{db}{U-d}), \\ U, & \text{if } x(0) = 0, \text{ and } t \in [a - \frac{1}{2} \frac{db}{U-d}, a); \\ \\ \dots\dots\dots \\ 0, & \text{if } x(0) \in (0, d(a - \frac{1}{2} \frac{db}{U-d})), \text{ and } t \in [0, \frac{x(0)}{d}), \\ d, & \text{if } x(0) \in (0, d(a - \frac{1}{2} \frac{db}{U-d})), \text{ and } t \in [\frac{x(0)}{d}, a - \frac{1}{2} \frac{db}{U-d}), \\ U, & \text{if } x(0) \in (0, d(a - \frac{1}{2} \frac{db}{U-d})), \text{ and } t \in [a - \frac{1}{2} \frac{db}{U-d}, a); \\ \\ \dots\dots\dots \\ 0, & \text{if } x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), \\ & \quad da + (U - \sqrt{U(U-d)})b), \text{ and } t \in [0, a - S_2(x(0))), \\ U, & \text{if } x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), \\ & \quad da + (U - \sqrt{U(U-d)})b), \text{ and } t \in [a - S_2(x(0)), a); \\ \\ \dots\dots\dots \\ 0, & \text{if } x(0) \geq da + (U - \sqrt{U(U-d)})b, \text{ and } t \in [0, a). \end{array} \right. \quad (4.184)
 \end{aligned}$$

We observe that during the first cycle, we have a *singular interval* (Type-1 Structure) for $x(0) \in (-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), d(a - \frac{1}{2} \frac{db}{U-d}))$, a *switching point* (Type-2 Structure) for $x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b]$, and Type-3 Structures for $x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d})$

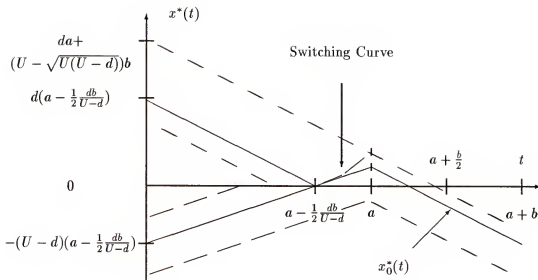


Figure 4.16. Optimal Trajectory for Problem P_2 .

(maximum production), and for $x(0) \geq da + (U - \sqrt{U(U-d)})b$ (no production). Figure 4.16 depicts some optimal trajectories. For an initial production surplus level $x(0)$ within the range $[-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), d(a - \frac{1}{2} \frac{db}{U-d})]$, the trajectory will reach the zero production surplus level at a time t where $t \in [0, a - \frac{1}{2} \frac{db}{U-d}]$. The trajectory will then stay at the zero production surplus level until time $t = a - \frac{1}{2} \frac{db}{U-d}$. For $t \in [a - \frac{1}{2} \frac{db}{U-d}, a]$, the machine is operated at maximum production rate before maintenance takes place during the interval $[a, a + b]$.

For an initial production surplus level $x(0)$ within the range $(d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b)$, the production rate initially is zero and the production surplus consequently decreases. At time $t_0 = a - S_2(x(0))$, the production rate switches from 0 to U . All the switching points form a switching curve $x_S(t_0)$ which is defined by the equation

$$S_2(x_S(t_0) + dt_0) = a - t_0, \quad \text{for } t_0 \in [a - C_r, a]. \quad (4.185)$$

Using Equation (4.174), we can solve for $x_S(t_0)$ and obtain

$$x_S(t_0) = Ub - \sqrt{Ub}\sqrt{U-d}\sqrt{b + 2(a - t_0)}, \quad \text{for } t_0 \in [a - \frac{1}{2} \frac{db}{U-d}, a]. \quad (4.186)$$

At the endpoints of the interval $[a - \frac{1}{2} \frac{db}{U-d}, a]$, we have

$$\begin{aligned}
 x_S(a - \frac{1}{2} \frac{db}{U-d}) &= Ub - \sqrt{Ub}\sqrt{U-d} \sqrt{b + 2(a - (a - \frac{1}{2} \frac{db}{U-d}))} \\
 &= Ub - \sqrt{Ub}\sqrt{U-d} \sqrt{b + \frac{db}{U-d}} \\
 &= Ub - \sqrt{Ub}\sqrt{U-d} \sqrt{b \frac{U}{U-d}} \\
 &= 0,
 \end{aligned}$$

and

$$\begin{aligned}
 x_S(a) &= Ub - \sqrt{Ub}\sqrt{U-d} \sqrt{b + 2(a-a)} \\
 &= Ub - \sqrt{Ub}\sqrt{U-d} \sqrt{b} \\
 &= b(U - \sqrt{U(U-d)}).
 \end{aligned}$$

Notice that the switching curve $x_S(t)$ for $t \in (a - \frac{1}{2} \frac{db}{U-d}, a]$ lies above the trajectory $x_0^*(t)$,

since we have $x_S(a - \frac{1}{2} \frac{db}{U-d}) = x_0^*(a - \frac{1}{2} \frac{db}{U-d}) = 0$ and

$$\begin{aligned}
 \frac{d}{dt} x_S(t) &= -\sqrt{Ub}\sqrt{U-d} \frac{1}{2\sqrt{b + 2(a-t_0)}} (-2) \\
 &= \frac{\sqrt{Ub}\sqrt{U-d}}{\sqrt{b + 2(a-t_0)}} \\
 &> \frac{\sqrt{Ub}\sqrt{U-d}}{\sqrt{b}}, \quad \text{since } t_0 \leq a \\
 &= \sqrt{U(U-d)} \\
 &> U-d \quad \text{for } t \in [a - \frac{1}{2} \frac{db}{U-d}, a] \\
 &= \frac{d}{dt} x_0^*(t),
 \end{aligned}$$

Thus, the optimal trajectory for an initial production surplus level $x(0) \in (d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b)$ does not touch the trajectory $x_0^*(t)$.

Since the optimal production policy $u_2^*(t)$ is determined by the initial production surplus level $x(0)$, we can therefore express the optimal performance measure in dependence of $x(0)$ as well. The optimal performance measure $J_2^*(x(0))$ satisfies

$$\begin{aligned}
J_2^*(x(0)) = & \left\{ \begin{aligned}
& \int_0^a \{x(0) + (U-d)t\}^2 dt + \int_a^{a+b} \{x(0) + Ua - dt\}^2 dt \\
& + J_1^*(x(0) + (U-d)a - db), \\
& \text{for } x(0) \leq -(U-d)(a - \frac{1}{2}\frac{db}{U-d}), \\
& \dots\dots\dots \\
& \int_0^{\frac{-x(0)}{U-d}} \{x(0) + (U-d)t\}^2 dt + \int_0^{\frac{1}{2}\frac{db}{U-d}} \{(U-d)t\}^2 dt \\
& + \int_0^b \{(U-d)\frac{1}{2}\frac{db}{U-d} - dt\}^2 dt + J_1^*(-\frac{1}{2}db), \\
& \text{for } x(0) \in [-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0], \\
& \dots\dots\dots \\
& \int_0^{\frac{1}{2}\frac{db}{U-d}} \{(U-d)t\}^2 dt + \int_0^b \{(U-d)\frac{1}{2}\frac{db}{U-d} - dt\}^2 dt + J_1^*(-\frac{1}{2}db), \\
& \text{for } x(0) = 0, \\
& \dots\dots\dots \\
& \int_0^{\frac{x(0)}{d}} \{x(0) - dt\}^2 dt + \int_0^{\frac{1}{2}\frac{db}{U-d}} \{(U-d)t\}^2 dt \\
& + \int_0^b \{(U-d)\frac{1}{2}\frac{db}{U-d} - dt\}^2 dt + J_1^*(-\frac{1}{2}db), \\
& \text{for } x(0) \in [0, d(a - \frac{1}{2}\frac{db}{U-d})], \\
& \dots\dots\dots \\
& \int_0^{a-S_2(x(0))} \{x(0) - dt\}^2 dt \\
& + \int_0^{S_2(x(0))} \{x(0) - d(a - S_2(x(0))) + (U-d)t\}^2 dt \\
& + \int_a^{a+b} \{x(0) + US_2(x(0)) - dt\}^2 dt \\
& + J_1^*(x(0) + US_2(x(0)) - d(a+b)), \\
& \text{for } x(0) \in [d(a - \frac{1}{2}\frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b], \\
& \dots\dots\dots \\
& \int_0^{a+b} \{x(0) - dt\}^2 dt + J_1^*(x(0) - d(a+b)), \\
& \text{for } x(0) \geq da + (U - \sqrt{U(U-d)})b.
\end{aligned} \right. \quad (4.187)
\end{aligned}$$

In Appendix C.8, we verify the following property:

$$\begin{aligned}\frac{d}{dx(0)}J_2^*(x_0) &< 0, & \text{for } x(0) < 0, \\ \frac{d}{dx(0)}J_2^*(x_0) &= 0, & \text{for } x(0) = 0, \\ \frac{d}{dx(0)}J_2^*(x_0) &> 0, & \text{for } x(0) > 0.\end{aligned}\tag{4.188}$$

From Equation (4.188), we conclude for arbitrary initial surplusses $x_1(0)$ and $x_2(0)$,

$$\begin{aligned}J_2^*(x_1(0)) &< J_2^*(x_2(0)) & \text{for } 0 \leq x_1(0) < x_2(0), \\ J_2^*(x_1(0)) &< J_2^*(x_2(0)) & \text{for } x_2(0) < x_1(0) \leq 0.\end{aligned}\tag{4.189}$$

According to Equation (4.187), we have

$$\begin{aligned}\frac{d}{dx(0)}J_2^*(x(0)) &= -\frac{1}{U-d}(x(0) + (U-d)(\frac{-x(0)}{U-d}))^2 + \int_0^{\frac{-x(0)}{U-d}} 2(x(0) + (U-d)t)dt \\ &= 2x(0)\frac{-x(0)}{U-d} + (U-d)(\frac{-x(0)}{U-d})^2 \\ &= -\frac{x^2(0)}{U-d}, \quad \text{for } x(0) \in [-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0].\end{aligned}\tag{4.190}$$

Also note that $J_2^*(x(0))$ has the following properties

$$\begin{aligned}\frac{d^2}{dx(0)^2}J_2^*(x(0)) &= \int_0^a 2dt + \int_a^{a+b} 2dt + \frac{d^2}{dx(0)^2}J_1^*(x(0) + (U-d)a - db) > 0, \\ &\text{for } x(0) \leq -(U-d)(a - \frac{1}{2}\frac{db}{U-d}),\end{aligned}\tag{4.191}$$

and

$$\frac{d^2}{dx(0)^2}J_2^*(x(0)) = \frac{-2x(0)}{U-d} > 0, \quad \text{for } x(0) \in (-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0).\tag{4.192}$$

We now show that $J_2^*(x(0))$ is strictly convex for $x(0) < 0$ by showing that the derivative of $J_2^*(x(0))$ is continuous for $x(0) = -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$. The continuity of $\frac{d}{dx(0)}J_2^*(x(0))$ at that point is established by showing that the left-hand-limit and the right-hand-limit are identical. The expression for the left-hand-limit, i.e., for an initial production surplus satisfying $x(0) < -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$, is obtained as follows,

$$\frac{d}{dx(0)}J_2^*(x(0)) = \int_0^a 2(x(0) + (U-d)t)dt + \int_a^{a+b} 2(x(0) + Ua - dt)dt$$

$$\begin{aligned}
& + \frac{d}{dx(0)} J_1^*(x(0) + (U-d)a - db) \\
= & 2x(0)a + (U-d)a^2 + 2x(0)b + 2Uab - d((a+b)^2 - a^2) \\
& + \frac{d}{dx(0)} J_1^*(x(0) + (U-d)a - db) \\
= & 2x(0)(a+b) + (U-d)a^2 + 2Uab - 2dab - db^2 \\
& + \frac{d}{dx(0)} J_1^*(x(0) + (U-d)a - db). \tag{4.193}
\end{aligned}$$

In order to complete the analysis of the left-hand-limit, we need to evaluate the expression

$\frac{d}{dx(0)} J_1^*(x(0) + (U-d)a - db)$. Notice that for $x(0) = -(U-d)(a - \frac{1}{2} \frac{db}{U-d})$, we have $x(0) + (U-d)a - db = -\frac{1}{2}db \in (-(U-d)(a - C_r), 0)$ and thus

$$\frac{d}{dx(0)} J_1^*(x(0) + (U-d)a - db) = -\frac{(-\frac{1}{2}db)^2}{U-d}, \quad \text{for } x(0) = -(U-d)(a - \frac{1}{2} \frac{db}{U-d}). \tag{4.194}$$

With Equations (4.193) and (4.194), we now verify that $\frac{d}{dx(0)} J_2^*(x(0))$ is continuous at

$$x(0) = -(U-d)(a - \frac{1}{2} \frac{db}{U-d}),$$

$$\begin{aligned}
& \lim_{x(0) \rightarrow -(U-d)(a - \frac{1}{2} \frac{db}{U-d}), x(0) < -(U-d)(a - \frac{1}{2} \frac{db}{U-d})} \frac{d}{dx(0)} J_2^*(x(0)) \\
= & -2(U-d)(a - \frac{1}{2} \frac{db}{U-d})(a+b) + (U-d)a^2 + 2(U-d)ab - db^2 - \frac{(-\frac{1}{2}db)^2}{U-d} \\
= & -2(U-d)a^2 - 2(U-d)ab + db(a+b) + (U-d)a^2 + 2(U-d)ab - db^2 - \frac{(-\frac{1}{2}db)^2}{U-d} \\
= & -(U-d)a^2 + dab - \frac{(-\frac{1}{2}db)^2}{U-d} \\
= & -\frac{(U-d)^2a^2 - (U-d)dab + (-\frac{1}{2}db)^2}{U-d} \\
= & -\frac{(-(U-d))^2(a^2 - 2a\frac{db}{U-d} + (-\frac{1}{2}\frac{db}{U-d})^2)}{U-d} \\
= & -\frac{(-(U-d))^2(a - \frac{db}{U-d})^2}{U-d} \\
= & -\frac{(-(U-d)(a - \frac{db}{U-d}))^2}{U-d} \\
= & \lim_{x(0) \rightarrow -(U-d)(a - \frac{1}{2} \frac{db}{U-d}), x(0) > -(U-d)(a - \frac{1}{2} \frac{db}{U-d})} \frac{d}{dx(0)} J_2^*(x(0)) \tag{4.195}
\end{aligned}$$

Therefore, we conclude that $J_2^*(x(0))$ is strictly convex for $x(0) < 0$.

This concludes our analysis of Problem P_2 . In the next section, we study the N -cycle problem.

4.5 Optimal solution to the N -cycle problem ($N > 3$)

In the previous section, we solved the two-cycle problem P_2 . In this section, we solve the N -cycle problem based on the solution of the $N - 1$ cycle problem for $N \geq 3$.

The time interval for the N -cycle problem is $[0, N(a + b))$. During the interval $[a + b, N(a + b))$, the optimal control can be determined based on the solution of Problem P_{N-1} . Further, $(a, a + b)$ is a maintenance interval. Thus, it remains to determine the optimal control during $[0, a]$ in order to solve the N -cycle problem.

The cost during the interval $[a, N(a + b)]$ is given by the expression $h_N(x(a))$:

$$h_N(x(a)) = h_1(x(a)) + J_{N-1}^*(x(a) - db).$$

The expression $h_N(x(a))$ is referred to as terminal cost at time $t = a$ and depends on the production surplus level at that time. It reflects the penalty during the interval $[a, N(a + b))$.

The function $h_N(x(a))$ satisfies

$$\frac{d}{dx(a)} h_N(x(a)) = 2x(a)b - db^2 + \frac{d}{dx(a)} J_{N-1}^*(x(a) - db). \quad (4.196)$$

The control problem for the N -cycle case can be restated as follows.

Problem P_N : Determine an admissible control $u^*(t)$, i.e., $u^*(t) \in [0, U]$, for $t \in [0, a]$, that causes the system

$$\frac{dx(t)}{dt} = u(t) - d$$

to follow an admissible trajectory $x^*(t)$ which minimizes the performance measure

$$J_N(u) = \int_0^a x^2(t) dt + h_N(x(a)).$$

The boundary conditions are $x(0) = x_0$ and $x(a)$ free.

Notice that the only difference between Problems P_1 , P_2 and P_N ($N \geq 3$) is the terminal cost. We will see that the analysis of Problem P_N is therefore similar to the analysis of Problems P_1 and P_2 . For example, the Hamiltonians for Problems P_1 , P_2 , and P_N are

independent of the terminal cost and are therefore identical. According to Pontryagin's Minimum Principle, the necessary optimality conditions for Problem P_N are as follows:

$$\frac{dx^*(t)}{dt} = u^*(t) - d, \quad (4.197)$$

$$\frac{dp^*(t)}{dt} = -2x^*(t), \quad (4.198)$$

$$\begin{aligned} x^{*2}(t) + p^*(t)[u^*(t) - d] &= \min_{u(t) \in [0, U]} \{x^{*2}(t) + p^*(t)[u(t) - d]\} \\ &= x^{*2}(t) - p^*(t)d + \min_{u(t) \in [0, U]} \{p^*(t)u(t)\}. \end{aligned} \quad (4.199)$$

The boundary condition for Problem P_N is

$$2x^*(a)b - db^2 - p^*(a) + \frac{d}{dx(a)} J_N^*(x(a) - db) = 0. \quad (4.200)$$

In addition, we have

$$x^{*2}(t) + p^*(t)[u^*(t) - d] = c_1, \quad \text{for } t \in [0, a], \quad (4.201)$$

and

$$u^*(t) = \begin{cases} 0, & \text{if } p^*(t) > 0, \\ U, & \text{if } p^*(t) < 0, \\ \beta, & \text{if } p^*(t) = 0. \end{cases} \quad (4.202)$$

where $\beta \in [0, U]$ is to be determined.

According to Equation (4.202), the optimal control law $u^*(t)$ is thus determined whenever we have $p^*(t) \neq 0$. In the case that roots exist for the function $p^*(t)$ within the interval $[0, a]$, in the same way as for Problem P_2 , we further distinguish between *singular intervals*, *switching times*, and *switching points*. The definitions of a singular interval and a switching point in Section 4.4 (see Equations (4.18) and (4.19)) apply here since we are considering the first cycle. The corresponding results regarding the possible control structures are summarized in the following.

4.5.1 Singular Intervals for Problem P_N

If there is a singular interval $[t_1, t_2] \subseteq [0, a]$, the optimal control satisfies

$$u^*(t) = \begin{cases} \beta_1, & \text{if } t \in [0, t_1), \\ d, & \text{if } t \in [t_1, t_2), \\ U, & \text{if } t \in [t_2, a), \end{cases} \quad (4.203)$$

where $\beta_1 \in \{0, U\}$.

4.5.2 Switching Points for Problem P_N

If $u^*(t)$ is an optimal control with a switching time $t_0 \in (0, a)$, necessary optimality conditions for the switching point and the initial production surplus level are

$$x^*(t_0) \geq 0, \quad (4.204)$$

$$x(0) > 0, \quad (4.205)$$

and it is further necessary that $u^*(t)$ satisfies

$$u^*(t) = \begin{cases} 0, & \text{if } t \in [0, t_0), \\ U, & \text{if } t \in [t_0, a). \end{cases} \quad (4.206)$$

4.5.3 Initial Control Level for Problem P_N

The initial control level $u^*(0)$ satisfies

$$u^*(0) = \begin{cases} 0, & \text{if } x(0) > 0, \\ d, & \text{if } x(0) = 0, \\ U, & \text{if } x(0) < 0. \end{cases} \quad (4.207)$$

4.5.4 Optimal Solution for Problem P_N

In this section, we solve for the optimal solution to Problem P_N . Based on the discussion above, the optimal control policy has three possible structures as defined in Section 4.4.4.

In order to facilitate the analysis, we partition the range of $x(0)$ in the following way:

$$D1: \quad x(0) = 0,$$

$$\begin{aligned}
\text{D2:} \quad & x(0) \in \left(- (U-d) \left(a - \frac{1}{2} \frac{db}{U-d}\right), 0\right), \\
\text{D3:} \quad & x(0) \leq - (U-d) \left(a - \frac{1}{2} \frac{db}{U-d}\right), \\
& \text{D3-1:} \quad x(0) \in \left[- (U-d)a, - (U-d) \left(a - \frac{1}{2} \frac{db}{U-d}\right)\right], \\
& \text{D3-2:} \quad x(0) < - (U-d)a, \\
\text{D4:} \quad & x(0) \geq d \left(a - \frac{1}{2} \frac{db}{U-d}\right), \\
& \text{D4-1:} \quad x(0) \geq d(a+b), \\
& \text{D4-2:} \quad x(0) \in \left[d \left(a - \frac{1}{2} \frac{db}{U-d}\right), d(a+b)\right], \\
\text{D5:} \quad & x(0) \in \left(0, d \left(a - \frac{1}{2} \frac{db}{U-d}\right)\right).
\end{aligned}$$

The derivation of the optimal control for Problem P_N for $N \geq 3$ is based on the following properties of Problem P_{N-1} .

1. *Property 1.* The first property concerns the derivative of the performance measure for a certain range of $x(0)$:

$$\frac{d}{dx(0)} J_{N-1}^*(x(0)) = -\frac{x^2(0)}{U-d}, \quad \text{for } x(0) \in \left[- (U-d) \left(a - \frac{1}{2} \frac{db}{U-d}\right), 0\right]. \quad (4.208)$$

For $N-1=2$, see Equation (4.190). Notice that Equation (4.208) implies

$$\begin{aligned}
\frac{d}{dC} J_{N-1}^*((U-d)C - db) &= (U-d) \left(-\frac{((U-d)C - db)^2}{U-d}\right) \\
&= -((U-d)C - db)^2 \\
&= -(U-d)^2 \left(C - \frac{db}{U-d}\right)^2, \quad (4.209)
\end{aligned}$$

for $(U-d)C - db \in \left[- (U-d) \left(a - \frac{1}{2} \frac{db}{U-d}\right), 0\right]$, or equivalently, $C \in \left[- \left(a - \frac{3}{2} \frac{db}{U-d}\right), \frac{db}{U-d}\right]$.

In particular, since $\frac{1}{2} \frac{db}{U-d} \in \left[- \left(a - \frac{3}{2} \frac{db}{U-d}\right), \frac{db}{U-d}\right]$, we have

$$\frac{d}{dC} J_{N-1}^*((U-d)C - db) = -\left(\frac{1}{2} db\right)^2, \quad \text{for } C = \frac{1}{2} \frac{db}{U-d}. \quad (4.210)$$

2. *Property 2.* The function $J_{N-1}^*(x(0))$ is strictly convex for $x(0) < 0$. For $N-1=2$, see Equations (4.191), (4.192), and (4.195).

3. Property 3.

$$\begin{aligned}
 \frac{d}{dx(0)} J_{N-1}^*(x_0) &< 0, & \text{for } x(0) < 0, \\
 \frac{d}{dx(0)} J_{N-1}^*(x_0) &= 0, & \text{for } x(0) = 0, \\
 \frac{d}{dx(0)} J_{N-1}^*(x_0) &> 0, & \text{for } x(0) > 0.
 \end{aligned} \tag{4.211}$$

Equivalently, we can write

$$\begin{aligned}
 J_{N-1}^*(x_1(0)) &< J_{N-1}^*(x_2(0)) & \text{if } 0 \leq x_1(0) < x_2(0), \\
 J_{N-1}^*(x_1(0)) &< J_{N-1}^*(x_2(0)) & \text{if } x_2(0) < x_1(0) \leq 0.
 \end{aligned} \tag{4.212}$$

We have shown that the three properties above are valid for $N = 2$. Assuming that they are true for arbitrary $N - 1$ ($N \geq 3$), we will show by induction in Section 4.5.4 that these properties are valid for general N .

Case D1: $x(0) = 0$

From Equation (4.207), we obtain for the initial control level $u^*(0) = d$ for $x(0) = 0$. With the same considerations as in Section 4.4.4, we conclude that the optimal production control has a Type-1 Structure with a singular interval $[0, a - C]$ where $C \in [0, a]$. The performance measure can be expressed in terms of C and satisfies

$$\begin{aligned}
 J_N(C) &= \int_0^a x^2(t) dt + h_N(x(a)) \\
 &= \int_0^C ((U - d)t)^2 dt + \int_0^b ((U - d)C - dt)^2 dt \\
 &\quad + J_{N-1}^*((U - d)C - db).
 \end{aligned} \tag{4.213}$$

We seek to determine the value of $C \in [0, a]$ which solves

$$\min\{J_N(C) : C \in [0, a]\}. \tag{4.214}$$

We differentiate $J_N(C)$ by using Leibniz' rule and obtain

$$\begin{aligned}
 \frac{d}{dC} J_N(C) &= ((U - d)C)^2 + \int_0^b 2(U - d)((U - d)C - dt) dt + \frac{d}{dC} J_{N-1}^*((U - d)C - db) \\
 &= (U - d)^2 \{C^2 + 2bC - \frac{d}{U - d} b^2\} + \frac{d}{dC} J_{N-1}^*((U - d)C - db).
 \end{aligned} \tag{4.215}$$

For $C = \frac{1}{2} \frac{db}{U-d}$, we have $\frac{d}{dC} J_{N-1}^*((U-d)C - db) = -(\frac{1}{2}db)^2$ according to Equation (4.210), and thus

$$\begin{aligned} \frac{d}{dC} J_N(C) &= (U-d)^2 \left\{ \left(\frac{1}{2} \frac{db}{U-d} \right)^2 + 2b \frac{1}{2} \frac{db}{U-d} - \frac{d}{U-d} b^2 \right\} - \left(\frac{1}{2} db \right)^2 \\ &= (U-d)^2 \left\{ 2b \frac{1}{2} \frac{db}{U-d} - \frac{d}{U-d} b^2 \right\} \\ &= 0, \quad \text{for } C = \frac{1}{2} \frac{db}{U-d}. \end{aligned} \quad (4.216)$$

The second derivative of $J_N(C)$ satisfies

$$\frac{d^2}{dC^2} J_N(C) = (U-d)^2 \{2C + 2b\} + \frac{d^2}{dC^2} J_{N-1}^*((U-d)C - db). \quad (4.217)$$

Since $J_{N-1}(x(0))$ is strictly convex for $x(0) < 0$, we have

$$\frac{d^2}{dC^2} J_{N-1}^*((U-d)C - db) > 0, \quad \text{for } C < \frac{db}{U-d}.$$

Thus, we conclude with Equation (4.217) that $J_N(C)$ is strictly convex for $C < \frac{db}{U-d}$, and thus, $C = \frac{1}{2} \frac{db}{U-d}$ solves (4.214). Therefore, for the first cycle, we have the same optimal control, optimal trajectory, and costate function as in Case C1 (see Equations (4.139), (4.140), and (4.141)).

Since we have $\frac{d}{dx(0)} J_{N-1}^*(x(0)) = -\frac{x^2(0)}{U-d}$ for $x(0) \in (-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), 0)$ according to Property 1, we conclude $\frac{d}{dx(0)} J_{N-1}^*(-\frac{1}{2}db) = -\frac{(\frac{1}{2}db)^2}{U-d}$. Since further the expressions for $x^*(a)$ and $p^*(a)$ are the same as in Case C1, the boundary condition (4.200) is verified in the same way as in Equation (4.142).

Optimality Condition

If $x_1^*(t)$ and $x_2^*(t)$ are two optimal trajectories for Problem P_N with initial production surplusses $x_1^*(0)$ and $x_2^*(0)$, respectively, then, the trajectories will not cross each other, i.e.,

$$\text{if } x_1^*(t') < x_2^*(t') \text{ for some } t' \in [0, a], \quad \text{then } x_1^*(t) \leq x_2^*(t) \text{ for all } t \in [0, a], \quad (4.218)$$

and in particular for $t' = 0$, we have

$$\text{if } x_1^*(0) < x_2^*(0), \quad \text{then} \quad x_1^*(t) \leq x_2^*(t), \quad \text{for } t \in [0, a]. \quad (4.219)$$

Case D2: $x(0) \in (-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0)$

From Equation (4.207), we have $u^*(0) = U$. Since the initial production surplus is negative, we can exclude a Type-2 Structure (switching point). Since the optimal trajectory $x^*(t)$ can not cross the trajectory $x_0^*(t)$, we conclude that it has a Type-1 Structure with singular interval $[\frac{-x(0)}{U-d}, a - \frac{1}{2}\frac{db}{U-d}]$. The optimal control, the optimal trajectory, and the costate function are thus governed by Equations (4.145), (4.146), and (4.148).

Case D3: $x(0) < -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$

First, since $x(0) < 0$, the initial control level satisfies $u^*(0) = U$ according to Equation (4.207). Therefore, a Type-2 Structure can be excluded, since it requires $u^*(0) = 0$. Thus, the optimal control has either a Type-1 Structure or a Type-3 Structure. That is, the optimal control either produces at maximum capacity or it has a singular interval. We further partition the initial surplus into two smaller regions, $x(0) \in [-(U-d)a, -(U-d)(a - \frac{1}{2}\frac{db}{U-d})]$ (Case D3-1) and $x(0) < -(U-d)a$ (Case D3-2), and show that the optimal control has a Type-3 Structure.

Case D3-1: $x(0) \in [-(U-d)a, -(U-d)(a - \frac{1}{2}\frac{db}{U-d})]$.

We either produce at *maximum capacity* (Type-3 Structure) or we have a *singular interval* $[t_1, t_2]$ (Type-1 Structure). As in Case C3-1, we can represent both control structures by

$$u(t) = \begin{cases} U, & \text{if } t \in [0, \frac{-x(0)}{U-d}), \\ d, & \text{if } t \in [\frac{-x(0)}{U-d}, a - C), \\ U, & \text{if } t \in [a - C, a), \end{cases} \quad (4.220)$$

where $C \in [0, a - \frac{-x(0)}{U-d}]$.

The performance measure satisfies

$$\begin{aligned} J_N(C) = & \int_0^{\frac{-x(0)}{U-d}} ((U-d)t)^2 dt + \int_0^C ((U-d)t)^2 dt + \int_0^b ((U-d)C - dt)^2 dt \\ & + J_{N-1}^*((U-d)C - db). \end{aligned} \quad (4.221)$$

We seek to determine the value of C which solves

$$\min\{J_N(C) : C \in [0, a - \frac{-x(0)}{U-d}]\}. \quad (4.222)$$

Since the first integral of the right-hand-side in Equation (4.221) is independent of C , the derivative $\frac{d}{dC}J_N(C)$ is identical to the derivative in Case D1. Thus, $J_N(C)$ is strictly convex for $C < \frac{db}{U-d}$ and has a unique minimum at $C = \frac{1}{2} \frac{db}{U-d}$. Since we have

$$a - \frac{-x(0)}{U-d} \leq \frac{1}{2} \frac{db}{U-d}, \quad \text{for } x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d}),$$

we conclude that $C = a - \frac{-x(0)}{U-d}$ solves (4.222). Equation (4.220) with $C = a - \frac{-x(0)}{U-d}$ leads to

$$u^*(t) = U, \quad \text{for } t \in [0, a]. \quad (4.223)$$

Case D3-2: $x(0) < -(U-d)a$.

Due to $x(0) < 0$, we have $u^*(0) = U$. Above, we have already excluded a Type-2 Structure. Here, we can exclude a Type-1 Structure with a singular interval as well, since $x(0) < -(U-d)a$ implies that no trajectory can reach the zero production surplus level for $t \in [0, a]$. Thus, we have a Type-3 Structure and produce at maximum capacity (see Equation (4.223)).

Thus, for the first cycle, we have the same control as in Case C3. The optimal control and the optimal trajectory are therefore governed according to Equations (4.154) and (4.155), respectively. With the boundary condition, we have for the costate function

$$p^*(t) = 2(a+b-t)x(0) + (U-d)(a^2 + 2ab - t^2) - db^2 + \frac{d}{dx(a)} J_{N-1}^*(x^*(a) - db). \quad (4.224)$$

This completes the analysis of Case D3 ($x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d})$).

Case D4: $x(0) > d(a - \frac{1}{2} \frac{db}{U-d})$.

Since the optimal trajectory can not cross the trajectory $x_0^*(t)$ according to Property (4.219) (which satisfies $x_0^*(t) > 0$ for $t \in (a - \frac{1}{2} \frac{db}{U-d}, a)$), we can exclude a singular interval (Type-1 Structure) from consideration. The production control for the remaining structures (Type-2 and Type-3) can be combined as in Equation (4.160). It results in the following

performance measure (see Equation (4.161)),

$$\begin{aligned} J_N(S) = & \int_0^{a-S} (x(0) - dt)^2 dt + \int_0^S (x(0) - d(a - S) + (U - d)t)^2 dt \\ & + \int_0^b (x(0) - da + US - dt)^2 dt + J_{N-1}^*(x(0) + US - d(a + b)). \end{aligned} \quad (4.225)$$

We seek to determine the solution of value of S which solves

$$\min\{J_N(S) : S \in [0, a]\}. \quad (4.226)$$

Its derivative is (see Equation (4.163))

$$\begin{aligned} \frac{d}{dS} J_N(S) = & U(U + d)\{S^2 + 2\frac{x(0) - da + Ub}{U + d}S - \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))\} \\ & + \frac{d}{dS} J_{N-1}^*(x(0) + US - d(a + b)). \end{aligned} \quad (4.227)$$

At this point, we further distinguish the cases $x(0) \geq d(a + b)$ (Case D4-1) and $x(0) \in [d(a - \frac{1}{2}\frac{db}{U-d}), d(a + b)]$ (Case D4-2).

Case D4-1: $x(0) \geq d(a + b)$.

According to Property 3, we have $\frac{d}{dS} J_{N-1}^*(x(0) + US - d(a + b)) \geq 0$ for $S \in [0, a]$ and $x(0) \geq d(a + b)$. Since the first term of the right-hand-side of Equation (4.227) is strictly positive, we have $\frac{d}{dS} J_N(S) > 0$ for $S \in [0, a]$ and $x(0) \geq d(a + b)$. This implies that the optimal control has a Type-3 Structure where the machine is shut down. That is, the optimal production control satisfies (see Equation (4.164))

$$u^*(t) = 0, \quad \text{for } t \in [0, a]. \quad (4.228)$$

The optimal trajectory satisfies (see Equation (4.165))

$$x^*(t) = x(0) - dt, \quad \text{for } t \in [0, a + b], \quad (4.229)$$

and the costate function is governed by (see Equation (4.166))

$$p^*(t) = 2x(0)(a + b - t) - d((a + b)^2 - t^2) + \frac{d}{dx(a)} J_{N-1}^*(x(0) - d(a + b)). \quad (4.230)$$

This completes the analysis of the case $x(0) \geq d(a + b)$.

Case D4-2: $x(0) \in [d(a - \frac{1}{2}\frac{db}{U-d}), d(a+b))$.

In this case, we also have Equation (4.227). We first consider the range of $x^*(a+b) = x(0) + US^* - d(a+b)$ where S^* solves (4.226). From Cases D1 and D4-1, we have

$$\begin{aligned} x^*(a+b) &= -\frac{1}{2}db, & \text{for } x(0) = 0, \\ x^*(a+b) &= 0, & \text{for } x(0) = d(a+b). \end{aligned}$$

Since the optimal trajectories will not cross each other (see the Optimality Condition (4.219)), we conclude

$$x^*(a+b) \in [-\frac{1}{2}db, 0], \quad \text{for } x(0) \in [0, d(a+b)]. \quad (4.231)$$

Since we have $-\frac{1}{2}db \geq -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$, we have further

$$x^*(a+b) \in [-\frac{1}{2}db, 0] \subseteq [-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0]. \quad (4.232)$$

According to Property 1, we have

$$\begin{aligned} \frac{d}{dx(0)} J_{N-1}^*(x(0) + US^* - d(a+b)) &= -\frac{x^{*2}(a+b)}{U-d}, \\ &\text{for } x^*(a+b) \in (-\frac{1}{2}db, 0). \end{aligned}$$

Therefore, Equation (4.227) can be written as

$$\begin{aligned} \frac{d}{dS} J_N(S) &= U(U+d) \{ S^2 + 2\frac{x(0) - da + Ub}{U+d} S - \frac{2b}{U+d} (d(a + \frac{b}{2}) - x(0)) \} \\ &+ -\frac{U}{U-d} (x(0) + US - d(a+b))^2. \end{aligned} \quad (4.233)$$

This expression is identical to the derivative in Case C4-2 (see Equation (4.172)). Therefore, the remaining analysis is identical to Case C4-2.

Case D5: $x(0) \in (0, d(a - \frac{1}{2}\frac{db}{U-d}))$

With the same reasoning as in Case C5, we conclude that we must have a singular interval (Type-1 Structure) (otherwise, the optimal trajectory for an initial production surplus $x(0) \in (0, d(a - \frac{1}{2}\frac{db}{U-d}))$ would cross the optimal trajectory for $x(0) = d(a - \frac{1}{2}\frac{db}{U-d})$).

The left endpoint t_1 of the singular interval satisfies $t_1 = \frac{x(0)}{d}$. Since this trajectory now meets the optimal trajectory $x_0^*(t)$, it coincides with $x_0^*(t)$ from then on and the right endpoint of the singular interval is $d(a - \frac{1}{2}\frac{db}{b-d})$. Thus, we have the same results as in Case C5 (see Equations (4.181), (4.182), and (4.183)).

Characterization of the optimal solution

Comparing the results for Problems P_2 and P_N , we obtain

$$u_N^*(t) = u_2^*(t), \quad \text{for } t \in [0, a]. \quad (4.234)$$

For the performance measure, we have

$$J_N^*(x(0)) = \left\{ \begin{array}{l} \int_0^a \{x(0) + (U-d)t\}^2 dt + \int_a^{a+b} \{x(0) + Ua - dt\}^2 dt \\ \quad + J_{N-1}^*(x(0) + (U-d)a - db), \\ \quad \text{for } x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d}), \\ \\ \dots\dots\dots \\ \int_0^{\frac{-x(0)}{U-d}} \{x(0) + (U-d)t\}^2 dt + \int_0^{\frac{1}{2} \frac{db}{U-d}} \{(U-d)t\}^2 dt \\ \quad + \int_0^b \{(U-d)\frac{1}{2} \frac{db}{U-d} - dt\}^2 dt + J_{N-1}^*(-\frac{1}{2}db), \\ \quad \text{for } x(0) \in [-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), 0], \\ \\ \dots\dots\dots \\ \int_0^{\frac{1}{2} \frac{db}{U-d}} \{(U-d)t\}^2 dt + \int_0^b \{(U-d)\frac{1}{2} \frac{db}{U-d} - dt\}^2 dt + J_{N-1}^*(-\frac{1}{2}db), \\ \quad \text{for } x(0) = 0, \\ \\ \dots\dots\dots \\ \int_0^{\frac{x(0)}{d}} \{x(0) - dt\}^2 dt + \int_0^{\frac{1}{2} \frac{db}{U-d}} \{(U-d)t\}^2 dt \\ \quad + \int_0^b \{(U-d)\frac{1}{2} \frac{db}{U-d} - dt\}^2 dt + J_{N-1}^*(-\frac{1}{2}db), \\ \quad \text{for } x(0) \in [0, d(a - \frac{1}{2} \frac{db}{U-d})], \\ \\ \dots\dots\dots \\ \int_0^{a-S_2(x(0))} \{x(0) - dt\}^2 dt \\ \quad + \int_0^{S_2(x(0))} \{x(0) - d(a - S_2(x(0))) + (U-d)t\}^2 dt \\ \quad + \int_a^{a+b} \{x(0) + US_2(x(0)) - dt\}^2 dt \\ \quad + J_{N-1}^*(x(0) + US_2(x(0)) - d(a+b)), \\ \quad \text{for } x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b], \\ \\ \dots\dots\dots \\ \int_0^{a+b} \{x(0) - dt\}^2 dt + J_{N-1}^*(x(0) - d(a+b)), \\ \quad \text{for } x(0) \geq da + (U - \sqrt{U(U-d)})b. \end{array} \right. \quad (4.235)$$

In order to complete our analysis, we need to show that the properties for Problem P_{N-1} which we listed at the beginning of this section also hold for Problem P_N .

1. *Property 1.* According to Equations (4.235), (4.187), and (4.190), we have

$$\frac{d}{dx(0)} J_N^*(x(0)) = -\frac{x^2(0)}{U-d}, \quad \text{for } x(0) \in [-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), 0].$$

2. *Property 2.* The function $J_N^*(x(0))$ is strictly convex for $x(0) < 0$. Using the strict convexity of $J_{N-1}(x(0))$ for $x(0) < 0$, this property is shown in the same way as it has been shown for Problem P_2 .

3. *Property 3.* From the Equations (4.187) and (4.235), we see

$$\frac{d}{dx(0)} J_N^*(x(0)) = \frac{d}{dx(0)} J_2^*(x(0)), \quad \text{for } x(0) \in [-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), d(a - \frac{1}{2} \frac{db}{U-d})]. \quad (4.236)$$

For $x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d})$, we have (see Equation (C.17))

$$\begin{aligned} \frac{d}{dx(0)} J_N^*(x(0)) &= \int_0^a 2(x(0) + (U-d)t)dt + \int_a^{a+b} 2(x(0) + Ua - dt)dt \\ &\quad + \frac{d}{dx(0)} J_{N-1}(x(0) + (U-d)a - db) \\ &= (U-d) \{ 2 \frac{x(0)}{U-d} (a+b) + (a+b)^2 - \frac{U}{U-d} b^2 \} \\ &\quad + \frac{d}{dx(0)} J_{N-1}(x(0) + (U-d)a - db). \end{aligned}$$

According to Equation (C.40), we have

$$(U-d) \{ 2 \frac{x(0)}{U-d} (a+b) + (a+b)^2 - \frac{U}{U-d} b^2 \} \leq 0, \quad \text{for } x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d}),$$

and since $x(0) + (U-d)a - db \leq -\frac{1}{2}db < 0$ for $x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d})$, we have according to Equation (4.211)

$$\frac{d}{dx(0)} J_{N-1}(x(0) + (U-d)a - db) < 0, \quad \text{for } x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d}).$$

In conclusion, we have

$$\frac{d}{dx(0)} J_N^*(x_0) < 0, \quad \text{for } x(0) \leq -(U-d)(a - \frac{1}{2} \frac{db}{U-d}).$$

For $x(0) \in [d(a - \frac{1}{2} \frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b]$ and $x(0) > da + (U - \sqrt{U(U-d)})b$, the result

$$\frac{d}{dx(0)} J_N(x(0)) > 0, \quad \text{for } x(0) \geq d(a - \frac{1}{2} \frac{db}{U-d})$$

is obtained in a similar way as for Problem P_2 (see Appendix C.8).

Thus, we have

$$\begin{aligned} \frac{d}{dx(0)} J_N^*(x_0) &< 0, & \text{for } x(0) < 0, \\ \frac{d}{dx(0)} J_N^*(x_0) &= 0, & \text{for } x(0) = 0, \\ \frac{d}{dx(0)} J_N^*(x_0) &> 0, & \text{for } x(0) > 0. \end{aligned} \tag{4.237}$$

This concludes the analysis of Problem P_N .

4.6 Transient State and Steady State

In this section, we analyze the *transient state* and the *steady state* of the system. First of all, we consider that the number of cycles N is sufficiently large, so that the analysis is not particularly dependent on the terminal conditions.

Figure 4.17 illustrates some trajectories with different initial surpluses. We observe that for $x(0) \in [-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), d(a - \frac{1}{2} \frac{db}{U-d})]$, all trajectories reach the point $x^*(a+b) = -\frac{1}{2}db$ at the end of the first cycle and will follow an identical pattern for the remaining up-down cycles from then on. This production surplus pattern is referred to as the *steady state trajectory*. In addition, for $x(0) \notin [-(U-d)(a - \frac{1}{2} \frac{db}{U-d}), d(a - \frac{1}{2} \frac{db}{U-d})]$, the trajectory contracts toward the zero production surplus level and will reach $x^*(n(a+b)) = -\frac{1}{2}db$ for the first time for some finite n (provided that either $x(0) \geq -\frac{1}{2}db$ or $(U-d)a > db$ holds). Then, it will follow the steady state trajectory.

We introduce the following definition for steady state: A system is said to be in steady state after n cycles if the surplus trajectory satisfies $x^*(n(a+b)) = -\frac{1}{2}db$ for the first time.

Therefore, the system is in transient state for $t \in [0, n(a+b)]$ and in steady state for $t > n(a+b)$ ($n < N$). Consequently, it would be of interest to determine n in terms of the initial production surplus.

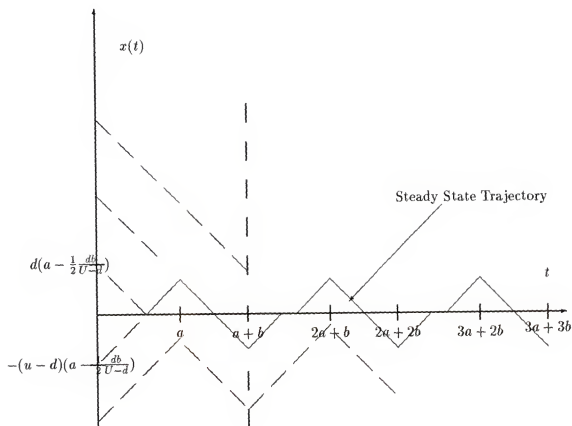


Figure 4.17. Transition and Steady State

The transition from an initial surplus state to the steady state can be described by the use of contraction mappings. Let's consider the first up-down cycle. Starting from the initial surplus $x(0)$, the optimal control policy will bring the production surplus to a level $x^*(a+b)$. Let $X(x(0))$ be the mapping function,

$$X(x(0)) = x^*(a+b).$$

Since the optimal control policy is known, the mapping function is well defined and is given by

$$X(x(0)) = \begin{cases} x(0) + (U-d)a - db, & \text{for } x(0) < -(U-d)(a - \frac{1}{2}\frac{db}{U-d}), \\ -\frac{1}{2}db, & \text{for } x(0) \in [-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), \\ & d(a - \frac{1}{2}\frac{db}{U-d})], \\ x(0) + US_2(x(0)) & \text{for } x(0) \in (d(a - \frac{1}{2}\frac{db}{U-d}), \\ & -d(a+b), \quad da + (U - \sqrt{U(U-d)})b), \\ x(0) - d(a+b), & \text{for } x(0) \geq da + (U - \sqrt{U(U-d)})b. \end{cases} \quad (4.238)$$

Subsequently, during the second cycle, the optimal control maps the production level $x^*(a+b) = X(x(0))$ to $x^*(2(a+b)) = X^2(x(0))$. After n cycles, we have

$$x^*(n(a+b)) = X^n(x(0)).$$

Consequently, the definition of steady state can be restated as follows: A system is said to have reached the steady state after n cycles if

$$\begin{aligned} X^{n-1}(x(0)) &\neq -\frac{1}{2}db, \\ \text{and } X^m(x(0)) &= -\frac{1}{2}db, \quad \text{for } n \leq m < N. \end{aligned} \quad (4.239)$$

According to the definition (4.239) and Equation (4.238), we now determine n for a given $x(0)$. Since we must have $x^*(n(a+b)) = -\frac{1}{2}db$ according to definition (4.239), $x(0)$ and n must satisfy

$$\begin{aligned} x(0) + n((U-d)a - db) &\geq -\frac{1}{2}db, & \text{if } x(0) < -\frac{1}{2}db, \\ x(0) - n(d(a+b)) &\leq -\frac{1}{2}db, & \text{if } x(0) \geq -\frac{1}{2}db. \end{aligned}$$

or equivalently,

$$x(0) \geq -n((U-d)a - db) - \frac{1}{2}db, \quad \text{if } x(0) < -\frac{1}{2}db, \quad (4.240)$$

$$x(0) \leq n(d(a+b)) - \frac{1}{2}db, \quad \text{if } x(0) \geq -\frac{1}{2}db. \quad (4.241)$$

Thus, we have

$$n = \begin{cases} \lceil \frac{x(0) + \frac{1}{2}db}{d(a+b)} \rceil, & \text{for } x(0) \geq -\frac{1}{2}db, \\ \lceil \frac{-(x(0) + \frac{1}{2}db)}{(U-d)a - db} \rceil, & \text{for } x(0) < -\frac{1}{2}db \text{ and } (U-d)a - db > 0. \end{cases} \quad (4.242)$$

The system does not reach the steady state for $x(0) < -\frac{1}{2}db$, if $(U-d)a - db = 0$. That is, the system has only marginal capacity and the surplus level can never reach $-\frac{1}{2}db$.

4.7 Summary

In this chapter, we study a production flow control problem with periodic maintenance. The system consists of a single machine and produces one product-type. The demand rate is constant. The objective is to trace the demand rate as closely as possible. Due to the maintenance disruption, the production deviates from the demand. An optimal control model is formulated to track the demand as closely as possible.

We have used Pontryagin's Minimum Principle to solve the one-cycle problem P_1 to optimality. We have discovered that the structure of the optimal control is dependent on the initial production surplus level $x(0)$. We have seen that there are three possible control structures which need to be considered for optimality. A Type-1 Structure includes a singular interval, a Type-2 Structure includes a switching point, and a Type-3 Structure has neither a singular interval nor a switching point.

The optimal control structure for the various ranges of $x(0)$ is largely governed by the optimal trajectory $x_0^*(t)$ for $x(0) = 0$. We have frequently used the Principle of Optimality in conjunction with the trajectory $x_0^*(t)$ to identify the optimal production control structure.

Next, we have used these techniques to solve the two-cycle and the N -cycle Problem. As already discovered for the one-cycle problem, the solution structure is different from a hedging point policy. With $N \geq 2$ cycles, the solution now has a symmetric trajectory $x_0^*(t)$ during the first maintenance interval $(a, a+b)$. The optimal control during the first cycle is the same as long as we have at least two cycles. The trajectory crosses the zero surplus level in the middle of the maintenance interval. We identified solutions both for the transient state and for the steady state. For a sufficiently large control horizon, starting from the

initial surplus, the trajectory contracts toward the zero surplus level and eventually reaches the steady state. In the steady state, the system follows an identical trajectory pattern for each of the remaining operation-maintenance cycles.

The study of this problem can be considered as a starting point for a new area of research. Currently, an extension of the model to a two-machine tandem line is being studied where the machines have the same maximum production capacity. Different machine capacities constitute a further extension. Systems with flexible machines would be another setting to apply Pontryagin's Minimum Principle as well as the case where the demand rate is a piecewise constant function.

APPENDIX A

PROOF OF LEMMA 1

Proof. For parts (a) and (b), we define ne , nt , Δ_1 , and Δ_2 as follows: Let ne be the number of jobs with $S_j < e$ and nt be the number of jobs with $C_j > d$. Further, define Δ_1 and Δ_2 by

$$\Delta_1 = \min(\min_j \{e - S_j | S_j < e\}, \min_j \{d - S_j | S_j < d\})$$

$$\Delta_2 = \min(\min_j \{C_j - e | C_j > e\}, \min_j \{C_j - d | C_j > d\})$$

- (a) For the decision variable case, let σ be a schedule which satisfies $C_j \neq d$ and $C_j \neq e$ for all j .

If $(ne - 1)\alpha > (nt)\beta$ ($(ne - 1)\alpha \leq (nt)\beta$), shift the due window $[e, d]$ to the left (right) by Δ_1 (Δ_2) units. This will decrease (increase) the earliness of $ne - 1$ jobs and increase (decrease) the tardiness of nt jobs and results in a schedule satisfying either $C_j = e$ or $C_j = d$ for some job J_j . The objective value of the new schedule is either maintained (in case of $(ne - 1)\alpha = (nt)\beta$) or improved by $\Delta_1((ne - 1)\alpha - (nt)\beta)$ ($\Delta_2((nt)\beta - (ne - 1)\alpha)$). Thus, an optimal schedule exists satisfying either $C_j = e$ or $C_j = d$ for some j . Please note that shifting the due window to the left is feasible ($e \geq 0$) due to the definition of Δ_1 and $S_{[1]} \geq 0$.

- (b) For the parameter case, assume that an optimal schedule σ satisfies neither (i), nor (ii), nor (iii). In this case, we construct another schedule by shifting the jobs:

If $(ne - 1)\alpha > (nt)\beta$ ($(ne - 1)\alpha \leq (nt)\beta$), shift all the jobs to the right (left) by Δ_1 ($\min(\Delta_2, S_{[1]})$) units. This will decrease (increase) the earliness of $ne - 1$ jobs and increase (decrease) the tardiness of nt jobs and results in a schedule satisfying either $C_j = e$ or $C_j = d$ for some job J_j or $S_{[1]} = 0$. The objective value of the new schedule is

either maintained (in case of $(ne-1)\alpha = (nt)\beta$) or improved by $\Delta_1((ne-1)\alpha - (nt)\beta)$ ($\min(\Delta_2, S_{[1]})((nt)\beta - (ne-1)\alpha)$). Thus, an optimal schedule exists satisfying either $C_j = e$ or $C_j = d$ for some j or $S_{[1]} = 0$. Please note that shifting the due window to the left is also feasible in this case.

- (c) (i) Let σ be an optimal schedule with $C_j = e$ for some job J_j . By *Property (CDW)2*, we know that the *E-jobs* are sequenced in LPT order.

Assume that there are two adjacent jobs J_j and J_k ($C_j = S_k$) with $p_j > p_k$ which both start after e , i.e., $S_j \geq e$. Construct schedule σ' by interchanging jobs J_j and J_k . We thus have $C'_j = C_k$ and $C'_k = C_j - (p_j - p_k)$ and therefore

$$\begin{aligned} \text{cost}(\sigma') - \text{cost}(\sigma) &= \beta((C'_k - d)^+) - (C_j - d)^+ \\ &= \beta((C_j - (p_j - p_k) - d)^+) - (C_j - d)^+ \leq 0 \end{aligned}$$

since $p_j - p_k > 0$. We repeat this process until all jobs J_l with $S_l \geq e$ are sequenced in SPT order. Since the cost does not increase, there is an optimal schedule where all jobs J_l with $S_l \geq e$ are sequenced in SPT order. Since the *E-jobs* are sequenced in LPT order, we have a *V-shape* around e .

- (ii) The proof in case of $C_j = d$ for some j is symmetric to the case $C_j = e$.
- (iii) Now we consider the parameter case where there is an optimal schedule σ with at least one job which is completed during the due window. Let job J_j be the job which satisfies both $e \leq C_j \leq d$ and $S_j < e$. By the same interchange argument as in the proof of part (c)(i), we can sequence jobs J_k with $S_k \geq C_j (\geq e)$ in SPT order and jobs J_k with $C_k \leq C_j (\leq d)$ in LPT order while either maintaining or improving the objective value. Thus, in this case ($S_{[1]} = 0$ and $e \leq C_j \leq d$ for some j), there is an optimal schedule where the jobs J_k with $S_k < e$ are sequenced in LPT order and the jobs J_k with $S_k \geq e$ are sequenced in SPT order.

QED

Lemma 4. Proof. We first demonstrate that an optimal solution exists satisfying

$$p_j \leq p_l \quad \text{for all } J_j \in W(\sigma) \text{ and } J_l \in E(\sigma) \cup T(\sigma). \quad (\text{A.1})$$

According to Lemma 1, parts (a), (c)(i) and (c)(ii), it is sufficient to consider the following two cases: Schedule σ satisfies either $C_k = d$ for some job J_k with a *V-shape* around d or $C_k = e$ for some job J_k with a *V-shape* around e .

1. $C_k = d$ for some job J_k with a *V-shape* around d :

Let job J_l be the *W-job* which starts at or before e and job J_j be the *T-job* which starts at d . Due to the *V-shape* structure around d , job J_l is the largest *W-job* and job J_j is the smallest *T-job*. Also, no *W-job* can be longer than an *E-job*. If schedule σ does not satisfy Equation (A.1), we must have $p_l > p_j$. We distinguish the cases $\sum_{k \in W(\sigma)} p_k - p_l + p_j \geq d - e$ and $\sum_{k \in W(\sigma)} p_k - p_l + p_j < d - e$:

(a) $\sum_{k \in W(\sigma)} p_k - p_l + p_j \geq d - e$:

In this case, obtain schedule σ' by interchanging jobs J_j and J_l such that $C_j(\sigma') = C_l(\sigma)$ holds (Figure A.1). We have

$$cw(J_l, \sigma) = cw(W, \sigma) = n(E(\sigma))\alpha = cw(W, \sigma') = cw(J_j, \sigma')$$

$$cw(J_j, \sigma) = n(T(\sigma))\beta = cw(J_l, \sigma')$$

$$cw(J_k, \sigma) = cw(J_k, \sigma') \quad \text{for } J_k \notin \{J_j, J_l\}$$

Applying Equation (2.7) of Lemma 3, we obtain

$$\begin{aligned} cost(\sigma') - cost(\sigma) &= (cw(J_j, \sigma') - cw(J_j, \sigma))p_j + (cw(J_l, \sigma') - cw(J_l, \sigma))p_l \\ &= (n(E(\sigma))\alpha - n(T(\sigma))\beta)(p_j - p_l) \leq 0, \end{aligned}$$

since $p_j < p_l$ and $n(E(\sigma))\alpha \geq n(T(\sigma))\beta$ by Lemma 2.

Obtain schedule σ'' from σ' by iteratively sequencing adjacent jobs starting before d in LPT order and the *T-jobs* in SPT order. Similarly to the proof in Lemma

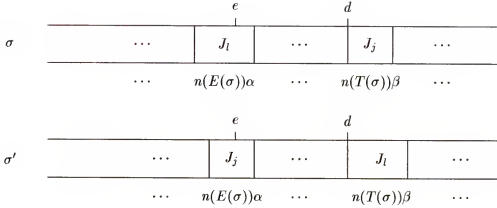


Figure A.1. $\sum_{k \in W(\sigma)} p_k - p_l + p_j \geq d - e$

1, parts (c)(i) and (c)(ii), we have $\text{cost}(\sigma'') \leq \text{cost}(\sigma')$. If $\max\{p_j : J_j \in W(\sigma'')\} \leq \min\{p_l : J_l \in T(\sigma'')\}$, schedule σ'' satisfies Equation (A.1) and we are done. Otherwise, we proceed as follows: If $\sum_{J_k \in W(\sigma'')} p_k - \max_{J_j \in W(\sigma'')} p_j + \min_{J_l \in T(\sigma'')} p_l \geq d - e$, we repeat the above process. After a finite number of iterations, we obtain a schedule σ^+ which either satisfies Equation (A.1) or $\sum_{J_k \in W(\sigma^+)} p_k - \max_{J_j \in W(\sigma^+)} p_j + \min_{J_l \in T(\sigma^+)} p_l < d - e$ holds. In the former case, we are finished, and in the latter case, we proceed with part (b).

(b) $\sum_{k \in W(\sigma)} p_k - p_l + p_j < d - e$:

In this case, obtain schedule σ' by interchanging jobs J_j and J_l such that $S_j(\sigma') = e$ holds (Figure A.2). We have

$$cw(J_k, \sigma) = cw(W, \sigma) = n(E(\sigma))\alpha \quad \text{for } J_k \in W(\sigma)$$

$$cw(J_k, \sigma') = cw(W, \sigma') = n(T(\sigma))\beta \quad \text{for } J_k \in W(\sigma)$$

$$cw(J_k, \sigma) = cw(J_k, \sigma') \quad \text{for } J_k \in E(\sigma) \cup T(\sigma)$$

Applying Equation (2.7) of Lemma 3, we obtain

$$\text{cost}(\sigma') - \text{cost}(\sigma) = \left\{ \sum_{J_k \in W(\sigma)} p_k - (d - e) \right\} \{n(T(\sigma))\beta - n(E(\sigma))\alpha\} \leq 0,$$

since $\sum_{J_k \in W(\sigma)} p_k \geq d - e$ and $n(T(\sigma))\beta \leq n(E(\sigma))\alpha$ according to Lemma 2.

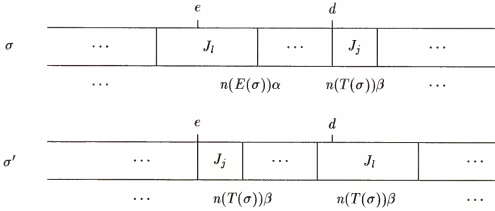


Figure A.2. $\sum_{k \in W(\sigma)} p_k - p_l + p_j < d - e$

Now obtain schedule σ'' by sequencing jobs with $C_k > e$ in SPT order. Notice that σ'' satisfies Equation (A.1) since $\max_{J_k \in W(\sigma'')} p_k \leq p_l = \min_{J_k \in E(\sigma)} p_k = \min_{J_k \in E(\sigma'')} p_k$.

2. $C_k = e$ for some job J_k with a *V-shape* around e :

This part is essentially symmetric to the case $C_k = d$. For the construction of schedule σ' , we interchange jobs J_j and J_l such that $S_j(\sigma') = S_l(\sigma)$. Also, schedule σ' needs to satisfy $S_{[1]} \geq 0$ after exchanging the two jobs. If necessary ($S_{[1]} < 0$), both the due window and the jobs need to be shifted to the right by $-S_{[1]} (> 0)$ which does not affect the objective function. The remainder is symmetric to the case $C_k = d$.

Finally, since the jobs are indexed in SPT order, by definition of $n w^*$ and Equation (A.1), we conclude that an optimal schedule σ exists such that $W(\sigma) = \{J_1, \dots, J_{n w^*}\}$ holds.

QED

APPENDIX B
ALGORITHM FOR $m = 2$ PARALLEL MACHINE CASE

Step 1a. Calculate Cumulative Weights:

$$(ne, nt, index) = (0, 0, n)$$

$$(E^1, E^2, T^1, T^2) = (\emptyset, \emptyset, \emptyset, \emptyset)$$

WHILE $index \geq nw$ DO

IF $(ne)\alpha \leq (nt + 1)\beta$ THEN

$$(cw_{index}, cw_{index-1}) = (ne\alpha, ne\alpha)$$

$$(E^1, E^2) \leftarrow (E^1 \cup \{index\}, E^2 \cup \{index - 1\})$$

$$index = index - 2$$

ELSE

$$(cw_{index}, cw_{index-1}) = ((nt + 1)\beta, (nt + 1)\beta)$$

$$(T^1, T^2) \leftarrow (T^1 \cup \{index\}, T^2 \cup \{index - 1\})$$

$$index = index - 2$$

Step 1b. Calculate Cost for E -jobs and T -jobs:

$$ET_{nw+3}^n = \sum_{j=nw+3}^n cw_j * p_j \qquad ET_{nw+2}^n = ET_{nw+3}^n + cw_{nw+2} * p_{nw+2}$$

Step 1c. Calculate A^{nw-1} and A^{nw} :

$$A^{nw-1} = \sum_{j=1}^{nw-1} p_j \qquad A^{nw} = A^{nw-1} + p_{nw}$$

Step 2a. Initialization of $cost_{min}^{nw}$:

$$cost_{min}^{nw} = \infty, \qquad t_{min}^{nw} \text{ and } Pos_{min}^{nw} \text{ need not be initialized.}$$

Step 2b. Using the results of the DP, examine all schedules with nw *WI-jobs*:

FOR ALL $J_k \in \{J_1, \dots, J_{nw+1}\}$ DO

$(Lower, Upper) = ((A^{nw+1} - p_k - (d - \epsilon))^+, \lfloor (A^{nw+1} - p_k)/2 \rfloor)$

(note the symmetry for $t > Upper$)

FOR $t = Lower$ TO $Upper$ DO

IF $f_k(nw + 1, t) = YES$ THEN

Examine both cases for job J_k (machine M^1 or M^2)

IF $t + p_k > d - \epsilon$ AND $(A^{nw+1} - p_k - t) + p_{nw+2} > d - \epsilon$

(i.e., $J_k \in JS^1$, $J_{nw+2} \in JS^2$, and $n(WI) = nw$) THEN

$$\begin{aligned}
 cost1 &= cw_{nw+1} * (t + p_k - (d - \epsilon)) \\
 &+ cw_{nw+2} * ((A^{nw+1} - p_k - t) + p_{nw+2} - (d - \epsilon)) \\
 &+ ET_{nw+3}^n \\
 cost2 &= cw_{nw+2} * (t + p_k - (d - \epsilon)) \\
 &+ cw_{nw+1} * ((A^{nw+1} - p_k - t) + p_{nw+2} - (d - \epsilon)) \\
 &+ ET_{nw+3}^n
 \end{aligned}$$

update $cost_{min}^{nw}$, t_{min}^{nw} , and position of J_k :

IF $cost1 < cost_{min}^{nw}$ THEN

$$cost_{min}^{nw} = cost1; \quad t_{min}^{nw} = t; \quad Pos_{min}^{nw} = M^1$$

IF $cost2 < cost_{min}^{nw}$ THEN

$$cost_{min}^{nw} = cost2; \quad t_{min}^{nw} = t; \quad Pos_{min}^{nw} = M^2$$

IF $t + p_{nw+2} > d - \epsilon$ AND $(A^{nw+1} - p_k) - t + p_k > d - \epsilon$

(i.e., $J_{nw+2} \in JS^1$, $J_k \in JS^2$, and $n(WI) = nw$) THEN

$$\begin{aligned}
 cost1 &= cw_{nw+1} * (t + p_{nw+2} - (d - \epsilon)) \\
 &+ cw_{nw+2} * (((A^{nw+1} - p_k) - t + p_k) - (d - \epsilon))
 \end{aligned}$$

$$\begin{aligned}
& + ET_{nw+3}^n \\
cost2 & = cw_{nw+2} * (t + p_{nw+2} - (d - \epsilon)) \\
& + cw_{nw+1} * (((A^{nw+1} - p_k) - t + p_k) - (d - \epsilon)) \\
& + ET_{nw+3}^n
\end{aligned}$$

update $cost_{min}^{nw}$, t_{min}^{nw} , and position of J_k :

IF $cost1 < cost_{min}^{nw}$ THEN

$$cost_{min}^{nw} = cost1; \quad t_{min}^{nw} = t; \quad Pos_{min}^{nw} = M^1$$

IF $cost2 < cost_{min}^{nw}$ THEN

$$cost_{min}^{nw} = cost2; \quad t_{min}^{nw} = t; \quad Pos_{min}^{nw} = M^2$$

Step 3. Using the results of the DP, examine all schedules with $nw - 1$ *WI-jobs*:

This is done similarly to Steps 2a and 2b.

Step 4. IF $cost_{min}^{nw} \leq cost_{min}^{nw-1}$ THEN

(the optimal schedule has nw *WI-jobs*, processes *WI-jobs* in machine M^1 with a cumulative processing time of t_{min}^{nw} , and positions job J_{nw+1} in machine Pos_{min}^{nw})

We now calculate the total processing time PE^i (PT^i) of the jobs in machine M^i which are processed before (after) the *WI-jobs* in machine M^i .

For the allocation of the jobs to machines M^1 and M^2 , we basically use the sets E^i and T^i which were generated in Step 1a. Yet we need to distinguish two cases $n - nw$ is even and $n - nw$ is odd. We also calculate minimum values ei_{min} and di_{min} for the location of the due window. The due window must satisfy $d \geq di_{min}$ and $d \geq ei_{min} + K$ for all $i \in \{1, 2\}$.

If $n - nw$ is even, the jobs J_{nw+1} and J_{nw+2} are allocated to the machines according to the calculation in Step 3.

If $n - nw$ is odd, we need to allocate jobs J_{nw+2} and J_{nw+3} at the same positions on different machines. First, we handle the case $n - nw$ is even.

IF $n - nw$ is even

FOR $i = 1$ TO 2 DO

$$PE^i = \sum_{j \in E^i \cap \{nw+3, \dots, n\}} p_j \quad PT^i = \sum_{j \in T^i \cap \{nw+3, \dots, n\}} p_j$$

IF $Pos_{min}^{nw} = M^1$ THEN

IF $(nw + 1) \in E^1 \cup E^2$

$$\text{THEN } PE^1 \leftarrow PE^1 + p_{nw+1} \quad d1_{min} = PE^1 + t_{min}^{nw}$$

$$\text{ELSE } PT^1 \leftarrow PT^1 + p_{nw+1} \quad e1_{min} = PE^1$$

IF $(nw + 2) \in E^1 \cup E^2$

$$\text{THEN } PE^2 \leftarrow PE^2 + p_{nw+2} \quad d2_{min} = PE^2 + (A^{nw} - t_{min}^{nw})$$

$$\text{ELSE } PT^2 \leftarrow PT^2 + p_{nw+2} \quad e2_{min} = PE^2$$

ELSE ($Pos_{min}^{nw} = M^2$)

IF $(nw + 1) \in E^1 \cup E^2$

$$\text{THEN } PE^2 \leftarrow PE^2 + p_{nw+1} \quad d2_{min} = PE^2 + (A^{nw} - t_{min}^{nw})$$

$$\text{ELSE } PT^2 \leftarrow PT^2 + p_{nw+1} \quad e2_{min} = PE^2$$

IF $(nw + 2) \in E^1 \cup E^2$

$$\text{THEN } PE^1 \leftarrow PE^1 + p_{nw+2} \quad d1_{min} = PE^1 + t_{min}^{nw}$$

$$\text{ELSE } PT^1 \leftarrow PT^1 + p_{nw+2} \quad e1_{min} = PE^1$$

The due window $[e, d]$ satisfies $e1_{min} \geq e$ and $d1_{min} \geq d$. Therefore, we have

$$d = \max(d1_{min}, d2_{min}, e1_{min} + K, e2_{min} + K)$$

This defines the location of the due window. The starting times of the jobs are obtained accordingly.

IF $n - nw$ is odd

In this case, the initialization of PE^i and PT^i starts from $nw + 4$:

FOR $i = 1$ TO 2 DO

$$PE^i = \sum_{j \in E^i \cap \{nw+4, \dots, n\}} p_j$$

$$PT^i = \sum_{j \in T^i \cap \{nw+4, \dots, n\}} p_j$$

The case distinctions are identical to the previous case. In addition, we take the following action: Whenever we allocated job J_{nw+2} in the previous case ($n - nw$ is even), we now ($n - nw$ is odd) allocate J_{nw+3} to the other machine.

Step 5. IF $cost_{min}^{nw} > cost_{min}^{nw-1}$ THEN

(the optimal schedule has $nw - 1$ *WI-jobs*, processes *WI-jobs* in machine M^1 with a cumulative processing time of t_{min}^{nw-1} , and positions job J_{nw} in machine Pos_{min}^{nw-1})

The procedure is similiar as in Step 4.

APPENDIX C

CALCULATION FOR A PRODUCTION SYSTEM WITH PERIODIC MAINTENANCE

C.1 The range of $S_1(x(0))$ for Problem P_1 (Case B4-2)

The goal of this section is to prove

$$S_1(x(0)) \leq C_r < a, \quad \text{for } x(0) \in [d(a - C_r), d(a + \frac{b}{2})], \quad (\text{C.1})$$

and

$$S_1(d(a - C_r)) = C_r < a. \quad (\text{C.2})$$

We proceed by showing that the function $S_1(x(0))$ is monotone decreasing,

$$\frac{d}{dx(0)} S_1(x(0)) < 0, \quad \text{for } x(0) \in (d(a - C_r), d(a + \frac{b}{2})),$$

and that the function value at $x(0) = d(a - C_r)$ is C_r ,

$$S_1(d(a - C_r)) = C_r.$$

Since we have

$$S_1(x(0)) = -\frac{x(0) - da + Ub}{U + d} + \sqrt{\left(\frac{x(0) - da + Ub}{U + d}\right)^2 + \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))}.$$

(see Equation (4.77)), we have for the first derivative

$$\begin{aligned} & \frac{d}{dx(0)} S_1(x(0)) \\ &= -\frac{1}{U + d} + \frac{2\left(\frac{x(0) - da + Ub}{U + d}\right)\frac{1}{U + d} - \frac{2b}{U + d}}{2\sqrt{\left(\frac{x(0) - da + Ub}{U + d}\right)^2 + \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))}} \\ &= \frac{1}{U + d} \left(\frac{\left(\frac{x(0) - da + Ub}{U + d}\right) - b}{\sqrt{\left(\frac{x(0) - da + Ub}{U + d}\right)^2 + \frac{2b}{U + d}(d(a + \frac{b}{2}) - x(0))}} - 1 \right) \\ &= \frac{1}{U + d} \left(\frac{(x(0) - d(a + b))}{\sqrt{(x(0) - da + Ub)^2 + 2b(U + d)(d(a + \frac{b}{2}) - x(0))}} - 1 \right). \end{aligned} \quad (\text{C.3})$$

Regrouping the radicant, we obtain

$$\begin{aligned}
& (x(0) - da + Ub)^2 + 2b(U + d)(d(a + \frac{b}{2}) - x(0)) \\
= & (x(0) - da)^2 + 2(x(0) - da)Ub + (Ub)^2 \\
& - 2(x(0) - da)Ub - 2(x(0) - da)db + (db)^2 + Udb^2 \\
= & (x(0) - da)^2 - 2(x(0) - da)db + (db)^2 + (Ub)^2 + Udb^2 \\
= & (x(0) - d(a + b))^2 + U(U + d)b^2 \\
> & (x(0) - d(a + b))^2.
\end{aligned}$$

Together with Equation (C.3), this yields

$$\frac{d}{dx(0)} S_1(x(0)) = \frac{1}{U + d} \left(\frac{(x(0) - d(a + b))}{\sqrt{(x(0) - d(a + b))^2 + U(U + d)b^2}} - 1 \right) \quad (\text{C.4})$$

$$< 0, \quad \text{for } x(0) \in [d(a - C_r), d(a + \frac{b}{2})]. \quad (\text{C.5})$$

It now remains to show

$$S(d(a - C_r)) = C_r. \quad (\text{C.6})$$

$$\begin{aligned}
& S(d(a - C_r)) - C_r \\
= & -\frac{Ub - dC_r}{U + d} + \sqrt{\left(\frac{Ub - dC_r}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) - C_r} \\
= & \sqrt{\left(\frac{Ub - dC_r}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) - \left(C_r + \frac{Ub - dC_r}{U + d}\right)} \\
= & \frac{\left(\frac{Ub - dC_r}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) - \left\{C_r + \frac{Ub - dC_r}{U + d}\right\}^2}{\sqrt{\left(\frac{Ub - dC_r}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) + \left(\frac{U}{U + d}C_r + \frac{Ub}{U + d}\right)}} \\
= & \frac{\frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) - C_r^2 - 2\frac{Ub - dC_r}{U + d}C_r}{\sqrt{\left(\frac{Ub - dC_r}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) + \left(\frac{U}{U + d}C_r + \frac{Ub}{U + d}\right)}} \\
= & \frac{(-1 + \frac{2d}{U + d})C_r^2 + (\frac{2bd}{U + d} - \frac{2Ub}{U + d})C_r + \frac{U - d}{U + d}\frac{db^2}{U - d}}{\sqrt{\left(\frac{Ub - dC_r}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) + \left(\frac{U}{U + d}C_r + \frac{Ub}{U + d}\right)}} \\
= & -\frac{U - d}{U + d} \frac{C_r^2 + 2bC_r - \frac{db^2}{U - d}}{\sqrt{\left(\frac{Ub - dC_r}{U + d}\right)^2 + \frac{2b}{U + d}\left(d\frac{b}{2} + dC_r\right) + \left(\frac{U}{U + d}C_r + \frac{Ub}{U + d}\right)}} \\
= & 0,
\end{aligned}$$

according to Equations (4.41) and (4.45).

In conclusion, we have verified Equation (C.1).

C.2 Switching Point for Problem P_1 (Case B4-2)

In this section, our goal is to verify

$$x^*(a - S_1(x(0))) \geq 0, \quad \text{for } x(0) \in (d(a - C_r), d(a + \frac{b}{2})) \quad (\text{C.7})$$

(compare Property (4.29)). Since the switching point satisfies

$$x^*(a - S_1(x(0))) = x(0) - d(a - S_1(x(0))),$$

we define a function $f(x(0))$ as follows,

$$f(x(0)) = x(0) - d(a - S_1(x(0))).$$

We verify Equation (C.7) by showing that the function is monotone increasing for the interval $(d(a - C_r), d(a + \frac{b}{2}))$ and has a root at $d(a - C_r)$. Therefore, we first show that the derivative of $f(x(0))$ with respect to $x(0)$ is positive. Since we have

$$\frac{d}{dx(0)} f(x(0)) = 1 + d \frac{d}{dx(0)} S_1(x(0)),$$

and

$$1 + d \frac{d}{dx(0)} S_1(x(0)) > 1 + U \frac{d}{dx(0)} S_1(x(0))$$

due to $\frac{d}{dx(0)} S_1(x(0)) < 0$ (see Equation (C.5)), it suffices to show

$$\frac{1}{U} + \frac{d}{dx(0)} S_1(x(0)) > 0. \quad (\text{C.8})$$

From Equation (C.4), we obtain

$$\begin{aligned} & \frac{d}{dx(0)} S_1(x(0)) + \frac{1}{U+d} \frac{U+d}{U} \\ = & \frac{1}{U+d} \left(\frac{(x(0) - d(a+b))}{\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2}} + \frac{d}{U} \right) \\ = & \frac{1}{U+d} \frac{d\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2} - U(d(a+b) - x(0))}{U\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2}} \\ = & \frac{1}{U+d} \frac{d^2\{(x(0) - d(a+b))^2 + U(U+d)b^2\} - U^2(d(a+b) - x(0))^2}{U\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2}} \\ & \quad + \frac{1}{d\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2} + U(d(a+b) - x(0))} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{U+d} \frac{U(U+d)d^2b^2 - (U^2 - d^2)(x(0) - d(a+b))^2}{U\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2}} \\
&\quad * \frac{1}{d\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2} + U(d(a+b) - x(0))} \\
&= \frac{U^2 - d^2}{U+d} \frac{-\{(x(0) - d(a+b))^2 - \frac{U}{U-d}d^2b^2\}}{U\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2}} \\
&\quad * \frac{1}{d\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2} + U(d(a+b) - x(0))} \\
&= (U-d) \frac{-\{(x(0) - d(a+b)) - d(C_r + b)\}\{(x(0) - d(a+b)) + d(C_r + b)\}}{U\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2}} \\
&\quad * \frac{1}{d\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2} + U(d(a+b) - x(0))} \\
&= (U-d) \frac{-\{(x(0) - d(a+2b+C_r))\}\{(x(0) - d(a-C_r))\}}{U\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2}} \\
&\quad * \frac{1}{d\sqrt{(x(0) - d(a+b))^2 + U(U+d)b^2} + U(d(a+b) - x(0))} \\
&> 0
\end{aligned}$$

for $x(0) \in (d(a-C_r), d(a+\frac{b}{2}))$. (C.9)

Thus, we have validated Equation (C.8),

$$\frac{d}{dx(0)}f(x(0)) > 0, \quad \text{for } x(0) \in (d(a-C_r), d(a+\frac{b}{2})).$$

Therefore, we have

$$\min\{f(x(0)) : x(0) \in (d(a-C_r), d(a+\frac{b}{2}))\} = f(d(a-C_r)).$$

By Equation (C.6), we have

$$f(d(a-C_r)) = d(a-C_r) - d(a-C_r) = 0.$$

This verifies Equation (C.7).

C.3 The Switching Curve

It has been observed in Section 4.3.4 that the optimal trajectories which have a switching point do not touch the trajectory $x_0^*(t)$ during the interval $[0, a]$ (see Equations (4.103) and (4.104)).

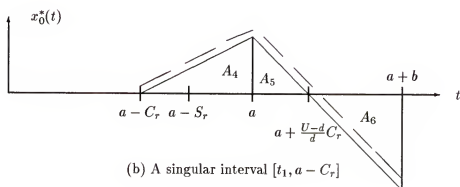
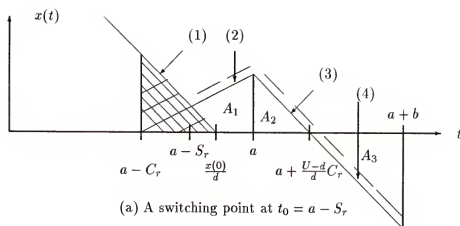


Figure C.1. Behavior of Type-2 Structures.

In this section, we explain this phenomenon from a different perspective. Figure C.1 shows two trajectories. In part (a) of Figure C.1, we have a trajectory for $x(0) \in (d(a - C_r), da)$ with a switching point at $a - S_r$ and in part (b) of Figure C.1, we have the Trajectory $x_0^*(t)$.

S_r in part (a) of Figure C.1 is given in such a way that the switching point lies on the Trajectory $x_0^*(t)$. The defining equation is

$$x(0) - d(a - S_r) = (U - d)((a - S_r) - (a - C_r)). \quad (C.10)$$

This results in

$$S_r = C_r - \frac{x(0) - d(a - C_r)}{U}. \quad (C.11)$$

We now express the performance measure for the trajectory in part (a) in terms of S_r and for the trajectory in part (b) in terms of C_r . The performance measure for the trajectory in part (a) is

$$\begin{aligned} J_1(S_r) &= \int_0^{a-S_r} (x(0) - dt)^2 dt + \int_{a-S_r}^a (x(0) - d(a - S_r) + (U - d)(t - (a - S_r)))^2 dt \\ &\quad + \int_a^{a+b} (x(0) - dt + US_r)^2 dt \\ &= \int_0^{a-S_r} (x(0) - dt)^2 dt + \int_{a-S_r}^a (x(0) - Ua + US_r + (U - d)t)^2 dt \\ &\quad + \int_a^{a+b} (x(0) - dt + US_r)^2 dt. \end{aligned}$$

We now want to demonstrate that the trajectory given in part (a) is not optimal. For this reason, we replace S_r by a parameter S , take the derivative, and evaluate the derivative for $S = S_r$. We have

$$\frac{d}{dS} J_1(S) = \int_{a-S}^a 2U(x(0) - Ua + US + (U - d)t) dt + \int_a^{a+b} 2U(x(0) - dt + US) dt.$$

We now rewrite the expressions $(x(0) - Ua + US + (U - d)t)$ and $(x(0) - dt + US)$. With $S = S_r = C_r - \frac{x(0) - d(a - C_r)}{U}$, we have

$$\begin{aligned} x(0) - Ua + US + (U - d)t &= x(0) - Ua + UC_r - x(0) + d(a - C_r) + (U - d)t \\ &= (U - d)(t - (a - C_r)), \end{aligned}$$

and

$$\begin{aligned} x(0) - dt + US &= x(0) - dt + UC_r - x(0) + d(a - C_r) \\ &= (U - d)C_r - d(t - a). \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} \frac{d}{dS} J_1(S_r) &= \int_{a-S_r}^a 2U(x(0) - Ua + US_r + (U - d)t)dt + \int_a^{a+b} 2U(x(0) - dt + US_r)dt \\ &= \int_{a-S_r}^a 2U(U - d)(t - (a - C_r))dt + \int_a^{a+b} 2U((U - d)C_r - d(t - a))dt \\ &= \int_{C_r-S_r}^{C_r} 2U(U - d)t dt + \int_0^b 2U((U - d)C_r - dt)dt \\ &= U(U - d)\{C_r^2 - (C_r - S_r)^2 + 2C_rb - db^2\} \\ &= -U(U - d)\{(C_r - S_r)^2\} \end{aligned} \tag{C.12}$$

$$< 0. \tag{C.13}$$

Thus, the trajectory in part (a) is not optimal and the optimal value S^* satisfies $S^* > S_r$ according to Equation (C.13). The optimal switching time $a - S^*$ is therefore smaller than $a - S_r$ and the resulting optimal trajectory will not touch the trajectory $x_0^*(t)$.

We come to the same conclusion when we consider parts (a) and (b) qualitatively. We know that Trajectory $x_0^*(t)$ is optimal for $x(0) = 0$. This implies $\frac{d}{dC} J_1(C_r) = 0$. We can partition the performance measure in two parts so that $J_1(C) = J_a(C) + J_b(C)$ where $J_a(C)$ refers to the interval $[0, a - S_r]$ and $J_b(C)$ refers to the interval $[a - S_r, a + b]$:

$$\begin{aligned} J_a(C) &= \int_{a-C}^{a-S_r} ((U - d)(t - (a - C)))^2 dt \\ &= \int_0^{C-S_r} (U - d)^2 t^2 dt \\ J_b(C) &= J_1(C) - J_a(C) \end{aligned} \tag{C.14}$$

According to Equation (C.14), it is not difficult to see that

$$\frac{d}{dC} J_a(C) = (U - d)^2 (C - S_r)^2 > 0. \tag{C.15}$$

Therefore, we have for $C = C_r$,

$$\frac{d}{dC} J_b(C_r) = \frac{d}{dC} (J_1(C_r) - J_a(C_r)) = 0 - \frac{d}{dC} J_a(C_r) < 0. \tag{C.16}$$

From parts (a) and (b) of Figure C.1, notice that the derivative of $J_b(C)$ for part (b) at $C = C_r$ is proportional to the derivative of $J_1(S)$ for part (a) at $S = S_r$. To see this, consider the areas A_1 to A_6 in Figure C.1. The triangles A_2 and A_5 as well as A_3 and A_6 correspond to each other. Polygon A_1 is contained in triangle A_4 . The time interval which is covered by A_1 , A_2 , and A_3 is $[a - S_r, a + b]$ and the time interval which is covered by A_4 , A_5 , and A_6 is $[a - C_r, a + b]$. Therefore, the change in the performance measure for part (b) for the interval $[a - S_r, a + b]$ ($\frac{d}{dC}J_b(C_r)$) is proportional to the change in performance measure for part (a) ($\frac{d}{dS}J_1(S_r)$).

Quantitatively, this observation is verified by Equations (C.12), (C.15), and (C.16):

$$(U - d)\frac{d}{dS}J_1(S_r) = -U\frac{d}{dC}J_a(C_r) = U\frac{d}{dC}J_b(C_r).$$

And since $\frac{d}{dC}J_b(C_r)$ is negative, $\frac{d}{dS}J_1(S_r)$ is negative as well. But this in turn implies that the parameter S must be increased and that the optimal value S^* must satisfy $S^* > S_r$. Consequently, the optimal trajectory for $x(0) > d(a - C_r)$ will not touch the Trajectory $x_0^*(t)$. We have thus shown that our quantitative conclusion can also be observed qualitatively by Figure C.1.

C.4 The First Derivative of $J_1^*(x(0))$

The first derivative of $J_1^*(x(0))$ with respect to $x(0)$ is calculated from Equation (4.105) according to the given ranges.

Case E1: $x(0) < -(U - d)(a - C_r)$.

$$\begin{aligned} \frac{d}{dx(0)}J_1^*(x(0)) &= \int_0^a 2(x(0) + (U - d)t)dt + \int_a^{a+b} 2(x(0) + Ua - dt)dt \\ &= 2x(0)a + (U - d)a^2 + 2x(0)b + 2Uab - d((a + b)^2 - a^2) \\ &= 2x(0)(a + b) + (U - d)(a^2 + 2ab + (b^2 - b^2)) - db^2 \\ &= (U - d)\left\{2\frac{x(0)}{U - d}(a + b) + (a + b)^2 - \frac{U}{U - d}b^2\right\}. \end{aligned} \quad (C.17)$$

Case E2: $x(0) \in (-(U - d)(a - C_r), 0)$.

$$\frac{d}{dx(0)}J_1^*(x(0)) = -\frac{1}{U - d}(x(0) + (U - d)\left\{\frac{-x(0)}{U - d}\right\})^2 + \int_0^{\frac{-x(0)}{U - d}} 2(x(0) + (U - d)t)dt$$

$$\begin{aligned}
&= 0 + 2x(0)\left\{\frac{-x(0)}{U-d}\right\} + (U-d)\left\{\frac{-x(0)}{U-d}\right\}^2 \\
&= -\frac{x^2(0)}{U-d}.
\end{aligned} \tag{C.18}$$

Case E3: $x(0) \in (0, d(a - C_r))$.

$$\begin{aligned}
\frac{d}{dx(0)}J_1^*(x(0)) &= \frac{1}{d}(x(0) - d\left\{\frac{x(0)}{d}\right\})^2 + \int_0^{\frac{x(0)}{d}} 2(x(0) - dt)dt \\
&= 0 + 2x(0)\left\{\frac{x(0)}{d}\right\} - d\left\{\frac{x(0)}{d}\right\}^2 \\
&= \frac{x^2(0)}{d}.
\end{aligned} \tag{C.19}$$

Case E4: $x(0) \in (d(a - C_r), d(a + \frac{b}{2}))$.

Due to the length of the calculation for this case, we calculate its derivative in Appendix

C.5. We have

$$\frac{d}{dx(0)}J_1^*(x(0)) = 2x(0)(a + b) - d(a + b)^2 + US_1^2(x(0)) + 2US_1(x(0))b. \tag{C.20}$$

Case E5: $x(0) > d(a + \frac{b}{2})$.

$$\begin{aligned}
\frac{d}{dx(0)}J_1^*(x(0)) &= \int_0^{a+b} 2(x(0) - dt)dt \\
&= 2x(0)(a + b) - d(a + b)^2.
\end{aligned} \tag{C.21}$$

This completes all cases for the range of $x(0)$. The result is summarized in Equation (4.106).

C.5 The Derivative of $J_1^*(x(0))$ (Case E4)

In this section, we calculate $\frac{d}{dx(0)}J_1^*(x(0))$ for $x(0) \in (d(a - C_r), d(a + \frac{b}{2}))$.

$$\begin{aligned}
&\frac{d}{dx(0)}J_1^*(x(0)) \\
&= \\
&\frac{d}{dx(0)}\left\{\int_0^{a-S_1(x(0))} (x(0) - dt)^2 dt \right. \\
&\quad + \int_0^{S_1(x(0))} (x(0) - d(a - S_1(x(0))) + (U - d)t)^2 dt \\
&\quad \left. + \int_a^{a+b} (x(0) - dt + US_1(x(0)))^2 dt\right\}
\end{aligned}$$

=

$$\begin{aligned}
& \left\{ -\frac{d}{dx(0)} S_1(x(0)) \right\} \{x(0) - d(a - S_1(x(0)))\}^2 \\
& + \int_0^{a-S_1(x(0))} 2(x(0) - dt) dt \\
& + \left\{ \frac{d}{dx(0)} S_1(x(0)) \right\} \{x(0) - d(a - S_1(x(0))) + (U - d)S_1(x(0))\}^2 \\
& + \int_0^{S_1(x(0))} 2\{x(0) - d(a - S_1(x(0))) + (U - d)t\} \left\{ 1 + d \frac{d}{dx(0)} S_1(x(0)) \right\} dt \\
& + \int_a^{a+b} 2\{x(0) - dt + U S_1(x(0))\} \\
& \quad \{1 + U \frac{d}{dx(0)} S_1(x(0))\} dt
\end{aligned}$$

=

$$\begin{aligned}
& \left\{ \int_0^{a-S_1(x(0))} 2(x(0) - dt) \quad dt \right. \\
& + \int_0^{S_1(x(0))} 2\{x(0) - d(a - S_1(x(0))) + (U - d)t\} \quad dt \\
& + \left. \int_a^{a+b} 2\{x(0) + U S_1(x(0)) - dt\} \quad dt \right\}
\end{aligned}$$

+

$$\begin{aligned}
& \frac{d}{dx(0)} S_1(x(0)) \\
& \quad \{-\{x(0) - d(a - S_1(x(0)))\}^2 \\
& \quad + \{x(0) - d(a - S_1(x(0))) + (U - d)S_1(x(0))\}^2 \\
& \quad + d\{2\{x(0) - d(a - S_1(x(0)))\}S_1(x(0))\} + (U - d)S_1^2(x(0))\} \\
& \quad + U\{2\{(x(0) + U S_1(x(0)))b\} - d((a + b)^2 - a^2)\} \quad \}. \tag{C.22}
\end{aligned}$$

We will now consider the factor which multiplies $\frac{d}{dx(0)} S_1(x(0))$ in Equation (C.22):

$$\begin{aligned}
& \{-\{x(0) - d(a - S_1(x(0)))\}^2 \\
& + \{x(0) - d(a - S_1(x(0))) + (U - d)S_1(x(0))\}^2 \\
& + d\{2\{x(0) - d(a - S_1(x(0)))\}S_1(x(0))\} + (U - d)S_1^2(x(0))\} \\
& + U\{2\{(x(0) + U S_1(x(0)))b\} - d((a + b)^2 - a^2)\} \quad \}
\end{aligned}$$

=

$$\begin{aligned}
& \{2\{x(0) - d(a - S_1(x(0)))\}\{(U - d)S_1(x(0))\} + \{(U - d)S_1(x(0))\}^2 \\
& + d\{2\{(x(0) - d(a - S_1(x(0))))S_1(x(0))\} + (U - d)S_1^2(x(0))\} \\
& + U\{2\{(x(0) + US_1(x(0)))b\} - d((a + b)^2 - a^2)\} \} \\
= & \\
& 2x(0)US_1(x(0)) - 2x(0)dS_1(x(0)) - 2aUdS_1(x(0)) + 2ad^2S_1(x(0)) \\
& + 2UdS_1^2(x(0)) - 2d^2S_1^2(x(0)) \\
& + U^2S_1^2(x(0)) - 2UdS_1^2(x(0)) + d^2S_1^2(x(0)) \\
& + 2x(0)dS_1(x(0)) - 2d^2aS_1(x(0)) + 2d^2S_1^2(x(0)) + UdS_1^2(x(0)) - d^2S_1^2(x(0)) \\
& + 2x(0)Ub + 2U^2bS_1(x(0)) - 2abUd - b^2Ud \\
= & \\
& \{2Ud - 2d^2 + U^2 - 2Ud + d^2 + 2d^2 + Ud - d^2\}S_1^2(x(0)) \\
& + \{2x(0)U - 2x(0)d - 2aUd + 2ad^2 + 2x(0)d - 2d^2a + 2U^2b\}S_1(x(0)) \\
& + \{2x(0)Ub - 2abUd - b^2Ud\} \\
= & \\
& \{U^2 + Ud\}S_1^2(x(0)) \\
& + \{2x(0)U - 2aUd + 2U^2b\}S_1(x(0)) \\
& + \{2x(0)Ub - 2abUd - b^2Ud\} \\
= & \\
& \{U(U + d)\}\{S_1^2(x(0)) + 2\frac{x(0) - da + Ub}{U + d}S_1(x(0)) + 2\frac{b}{U + d}(x(0) - d(a + \frac{b}{2}))\} \\
= & 0, \tag{C.23}
\end{aligned}$$

according to Equations (4.41) and (4.77). Thus, substituting the result of Equation (C.23) into Equation (C.22), we have

$$\begin{aligned}
\frac{d}{dx(0)}J_1^*(x(0)) = & \\
& \{ \int_0^{a-S_1(x(0))} 2(x(0) - dt) \quad dt
\end{aligned}$$

$$\begin{aligned}
& + \int_0^{S_1(x(0))} 2\{x(0) - d(a - S_1(x(0))) + (U - d)t\} dt \\
& + \int_a^{a+b} 2\{x(0) + U S_1(x(0)) - dt\} dt \\
& = \\
& \{ \int_0^{a-S_1(x(0))} 2(x(0) - dt) dt \\
& + \int_{a-S_1(x(0))}^a 2\{x(0) - dt + U(t - (a - S_1(x(0))))\} dt \\
& + \int_a^{a+b} 2\{x(0) - dt + U S_1(x(0))\} dt \} \\
& = \\
& \int_0^{a+b} 2(x(0) - dt) dt \\
& + \int_0^{S_1(x(0))} 2U t dt \\
& + \int_a^{a+b} 2U S_1(x(0)) dt \} \\
& = \\
& 2x(0)(a+b) - d(a+b)^2 + U S_1^2(x(0)) + 2U S_1(x(0))b. \quad (C.24)
\end{aligned}$$

C.6 The Second Derivative of $J_1^*(x(0))$

According to Equation (4.106) (see also Cases E1 to E5 in Appendix C.4), we obtain the following expressions for the second derivative of $J_1^*(x(0))$.

Case F1: $x(0) < -(U - d)(a - C_r)$.

In this case, we have from Equation (4.105) according to Leibniz' rule

$$\begin{aligned}
\frac{d^2}{dx^2(0)} J_1^*(x(0)) &= \frac{d}{dx(0)} \left\{ \int_0^a 2(x(0) + (U - d)t) dt + \int_a^{a+b} 2(x(0) + Ua - dt) dt \right\} \\
&= \int_0^a 2 dt + \int_a^{a+b} 2 dt \\
&= 2(a + b). \quad (C.25)
\end{aligned}$$

Case F2: $x(0) \in (-(U - d)(a - C_r), 0)$.

According to Equation (4.105), we derive

$$\frac{d^2}{dx^2(0)} J_1^*(x(0)) = \frac{d}{dx(0)} \left\{ -\frac{1}{U - d} (x(0) + (U - d) \left\{ \frac{-x(0)}{U - d} \right\})^2 \right\}$$

$$\begin{aligned}
& + \int_0^{\frac{-x(0)}{U-d}} 2(x(0) + (U-d)t)dt \} \\
& = 0 - \frac{1}{U-d} 2(x(0) + (U-d)\frac{-x(0)}{U-d}) + \int_0^{\frac{-x(0)}{U-d}} 2dt \\
& = \frac{-2x(0)}{U-d}.
\end{aligned} \tag{C.26}$$

Case F3: $x(0) \in (0, d(a - C_r))$.

From Equation (4.105), we derive

$$\begin{aligned}
\frac{d^2}{dx^2(0)} J_1^*(x(0)) &= \frac{d}{dx(0)} \left\{ \frac{1}{d} (x(0) - d\frac{x(0)}{d})^2 + \int_0^{\frac{x(0)}{d}} 2(x(0) - dt)dt \right\} \\
&= 0 + \frac{1}{d} 2(x(0) - d(\frac{x(0)}{d})) + \int_0^{\frac{x(0)}{d}} 2dt \\
&= \frac{2x(0)}{d}.
\end{aligned} \tag{C.27}$$

Case F4: $x(0) \in (d(a - C_r), d(a + \frac{b}{2}))$.

In Appendix C.5, we derive from Equation (4.105)

$$\begin{aligned}
& \frac{d^2}{dx^2(0)} J_1^*(x(0)) \\
&= \frac{2U}{U+d} \left\{ a - S_1(x(0)) - x(0) \frac{d}{dx(0)} S_1(x(0)) \right. \\
& \quad \left. + d(a+b) \left(\frac{1}{U} + \frac{d}{dx(0)} S_1(x(0)) \right) \right\}.
\end{aligned} \tag{C.28}$$

Case F5: $x(0) \geq d(a + \frac{b}{2})$.

In this case, we have

$$\begin{aligned}
\frac{d^2}{dx^2(0)} J_1^*(x(0)) &= \frac{d}{dx(0)} \left\{ \int_0^{a+b} 2(x(0) - dt)dt \right\} \\
&= \int_0^{a+b} 2dt \\
&= 2(a+b).
\end{aligned} \tag{C.29}$$

This completes all cases for the range of $x(0)$. The result is summarized in Equation (4.109).

C.7 A Property of $\frac{d}{dx(0)}J_1^*(x(0))$

In this subsection, we want to validate Equation (4.107). That is, we want to show

$$\begin{aligned}\frac{d}{dx(0)}J_1^*(x_0) &< 0, & \text{for } x(0) < 0, \\ \frac{d}{dx(0)}J_1^*(x_0) &= 0, & \text{for } x(0) = 0, \\ \frac{d}{dx(0)}J_1^*(x_0) &> 0, & \text{for } x(0) > 0.\end{aligned}$$

We proceed by considering the ranges for $x(0)$ according to Cases C1 to C5 (see Equation (4.105)).

Case G1: $x(0) < -(U-d)(a-C_r)$.

According to Equation (C.17), we have

$$\frac{d}{dx(0)}J_1^*(x(0)) = (U-d)\left\{2\frac{x(0)}{U-d}(a+b) + (a+b)^2 - \frac{U}{U-d}b^2\right\}, \quad (\text{C.30})$$

and conclude

$$\sup\left\{\frac{d}{dx(0)}J_1^*(x(0)) : x(0) < -(U-d)(a-C_r)\right\} = \frac{d}{dx(0)}J_1^*(-(U-d)(a-C_r)).$$

We therefore have for the left-hand-limit of $\frac{d}{dx(0)}J_1^*(x(0))$ at $x(0) = -(U-d)(a-C_r)$

$$\begin{aligned}& \lim_{x(0) \rightarrow -(U-d)(a-C_r), x(0) < -(U-d)(a-C_r)} \frac{d}{dx(0)}J_1^*(x(0)) \\&= (U-d)\left\{-2(a-C_r)(a+b) + (a+b)^2 - \frac{U}{U-d}b^2\right\} \\&= (U-d)\left\{-2(a+b - \sqrt{\frac{U}{U-d}b})(a+b) + (a+b)^2 - \frac{U}{U-d}b^2\right\} \\&= (U-d)\left\{-(a+b)^2 + 2\sqrt{\frac{U}{U-d}b}(a+b) - \frac{U}{U-d}b^2\right\} \\&= -(U-d)\left\{((a+b) - \sqrt{\frac{U}{U-d}b})^2\right\} \\&= -(U-d)\{a-C_r\}^2 \\&< 0,\end{aligned}$$

and conclude

$$\frac{d}{dx(0)}J_1^*(x(0)) < 0, \quad \text{for } x(0) < -(U-d)(a-C_r).$$

Case G2: $x(0) \in -(U - d)(a - C_r), 0$.

According to Equation (C.18), we have

$$\frac{d}{dx(0)} J_1^*(x(0)) = -\frac{x^2(0)}{U - d}, \quad (\text{C.31})$$

and conclude

$$\frac{d}{dx(0)} J_1^*(x(0)) < 0, \quad \text{for } x(0) \in -(U - d)(a - C_r), 0.$$

Case G3: $x(0) \in (0, d(a - C_r))$.

According to Equation (C.19), we have

$$\frac{d}{dx(0)} J_1^*(x(0)) = \frac{x^2(0)}{d}, \quad (\text{C.32})$$

and conclude

$$\frac{d}{dx(0)} J_1^*(x(0)) > 0, \quad \text{for } x(0) \in (0, d(a - C_r)).$$

Case G4: $x(0) \in (d(a - C_r), d(a + \frac{b}{2}))$.

In the same way as in Cases G1, G2, and G3, it can be shown that the derivative of $J_1^*(x(0))$ with respect to $x(0)$ according to Equation (4.105) is positive for the range $(d(a - C_r), d(a + \frac{b}{2}))$. It turns out that this requires tedious calculation.

Instead, we show that

$$J_1^*(x(0) + \Delta x(0)) - J_1^*(x(0)) > 0, \\ \text{for } \Delta x(0) \in (0, d(a - C_r)) \text{ and } x(0) \in (d(a - C_r), d(a + \frac{b}{2}) - \Delta x(0)) \quad (\text{C.33})$$

For simplicity, we denote $x(0)$ by $x_1(0)$ and $x(0) + \Delta x(0)$ by $x_2(0)$. According to Equation (C.1), we have

$$S_1(x(0)) \leq C_r < a, \quad \text{for } x(0) \in [d(a - C_r), d(a + \frac{b}{2})].$$

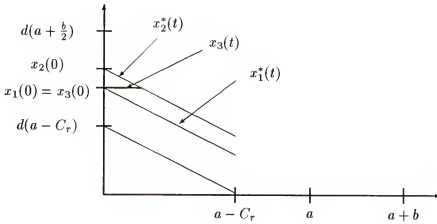


Figure C.2. $J_1^*(x(0))$ is monotone increasing for $x(0) \in (d(a - C_r), d(a + \frac{b}{2}))$

Since $x_1(0), x_2(0) \in (d(a - C_r), d(a + \frac{b}{2}))$, we have according to Equation (4.85)

$$\begin{aligned} u_1^*(t) &= 0, & \text{for } t \in [0, a - C_r] \subseteq [0, a - S_1(x_1(0))], \\ u_2^*(t) &= 0, & \text{for } t \in [0, a - C_r] \subseteq [0, a - S_1(x_2(0))]. \end{aligned}$$

The optimal trajectories $x_1^*(t)$ and $x_2^*(t)$ during $[0, a - C_r]$ are shown in Figure C.2. In order to compare $J_1^*(x_1(0))$ and $J_1^*(x_2(0))$, we consider a third trajectory $x_3(t)$ which is obtained as follows: We start with an initial production surplus level of $x_3(0) = x_1(0)$ and use the production control

$$u_3(t) = \begin{cases} d & \text{for } t \in [0, \frac{x_2(0) - x_1(0)}{d}), \\ u_2^*(t) & \text{for } t \in [\frac{x_2(0) - x_1(0)}{d}, a), \end{cases}$$

which results in a trajectory $x_3(t)$ with

$$x_3(t) = \begin{cases} x_3(0) & \text{for } t \in [0, \frac{x_2(0) - x_1(0)}{d}), \\ x_2^*(t) & \text{for } t \in [\frac{x_2(0) - x_1(0)}{d}, a), \end{cases}$$

Thus, the trajectory $x_3(t)$ starts at the same initial production surplus as trajectory $x_1^*(t)$, stays on this level until it reaches trajectory $x_2^*(t)$ and during $[\frac{x_2(0) - x_1(0)}{d}, a + b]$ is identical to trajectory $x_2^*(t)$. Since the trajectory $x_3(t)$ is positive and below trajectory

$x_2^*(t)$ during $[0, \frac{x_2(0)-x_1(0)}{d}]$ and subsequently coincides with it, we conclude

$$\int_0^T x_3^2(t)dt \leq J_1^*(x_2^*(t)). \quad (C.34)$$

Since trajectory $x_1^*(t)$ is the optimal trajectory starting from $x_1(0) = x_3(0)$, we have

$$J_1^*(x_1^*(t)) < \int_0^T x_3^2(t)dt. \quad (C.35)$$

From Equations (C.34) and (C.35), we conclude

$$J_1^*(x_1(0)) < J_1^*(x_2(0)).$$

Consequently, we have

$$\frac{d}{dx(0)} J_1^*(x(0)) = \lim_{\Delta x(0) \rightarrow 0} \frac{J_1^*(x(0) + \Delta x(0)) - J_1^*(x(0))}{\Delta x(0)} > 0, \quad (C.36)$$

for $x(0) \in (d(a - C_r), d(a + \frac{b}{2}))$. Thus, we have shown in a different way that the derivative for this range of initial production surplus values is positive.

Case G5: $x(0) > d(a + \frac{b}{2})$.

According to Equation (C.21), we have for $x(0) > d(a + \frac{b}{2})$

$$\frac{d}{dx(0)} J_1^*(x(0)) = 2x(0)(a + b) - d(a + b)^2. \quad (C.37)$$

We conclude for $x(0) \geq d(a + \frac{b}{2})$,

$$\begin{aligned} \frac{d}{dx(0)} J_1^*(x(0)) &\geq 2d(a + \frac{b}{2})(a + b) - d(a + b)^2, \\ &= d(2a + b)(a + b) - d(a + b)^2, \\ &= da(a + b), \\ &> 0. \end{aligned} \quad (C.38)$$

Finally, we observe

$$\lim_{x(0) \rightarrow 0, x(0) > 0} \frac{d}{dx(0)} J_1^*(x(0)) = \lim_{x(0) \rightarrow 0, x(0) < 0} \frac{d}{dx(0)} J_1^*(x(0)) = 0.$$

In summary, we have validated Equation (4.107)

$$\begin{aligned}\frac{d}{dx(0)}J_1^*(x_0) &< 0, & \text{for } x(0) < 0, \\ \frac{d}{dx(0)}J_1^*(x_0) &= 0, & \text{for } x(0) = 0, \\ \frac{d}{dx(0)}J_1^*(x_0) &> 0, & \text{for } x(0) > 0.\end{aligned}$$

C.8 A Property of $\frac{d}{dx(0)}J_2^*(x(0))$

In this subsection, we want to validate Equation (4.188). That is, we want to show

$$\begin{aligned}\frac{d}{dx(0)}J_2^*(x_0) &< 0, & \text{for } x(0) < 0, \\ \frac{d}{dx(0)}J_2^*(x_0) &= 0, & \text{for } x(0) = 0, \\ \frac{d}{dx(0)}J_2^*(x_0) &> 0, & \text{for } x(0) > 0.\end{aligned}$$

We proceed by considering the ranges for $x(0)$ according to Equation (4.187).

Case H1: $x(0) < -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$.

According to Equation (4.187), we derive in the same way as in Equation (C.30),

$$\begin{aligned}\frac{d}{dx(0)}J_2^*(x(0)) &= (U-d)\{2\frac{x(0)}{U-d}(a+b) + (a+b)^2 - \frac{U}{U-d}b^2\} \\ &\quad + \frac{d}{dx(0)}J_1^*(x(0) + (U-d)a - db),\end{aligned}\tag{C.39}$$

and conclude

$$\sup\{\frac{d}{dx(0)}J_2^*(x(0)) : x(0) < -(U-d)(a - \frac{1}{2}\frac{db}{U-d})\} = \frac{d}{dx(0)}J_2^*(-(U-d)(a - \frac{1}{2}\frac{db}{U-d})).$$

We have $\frac{d}{dx(0)}J_1^*(x(0) + (U-d)a - db) < 0$ for $x(0) = -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$, since $x(0) + (U-d)a - db = -\frac{1}{2}db < 0$ for $x(0) = -(U-d)(a - \frac{1}{2}\frac{db}{U-d})$. We conclude

$$\begin{aligned}&\frac{d}{dx(0)}J_2^*(-(U-d)(a - \frac{1}{2}\frac{db}{U-d})) \\ &< (U-d)\{-2(a - \frac{1}{2}\frac{db}{U-d})(a+b) + (a+b)^2 - \frac{U}{U-d}b^2\} \\ &= (U-d)\{(a+b)(a+b - 2(a - \frac{1}{2}\frac{db}{U-d})) - \frac{U}{U-d}b^2\} \\ &= (U-d)\{(a+b)(-a+b + \frac{db}{U-d}) - \frac{U}{U-d}b^2\} \\ &= (U-d)\{-a^2 + b^2 + \frac{dab}{U-d} + \frac{db^2}{U-d} - \frac{Ub^2}{U-d}\}\end{aligned}$$

$$\begin{aligned}
&= (U-d)\{-a^2 + \frac{dab}{U-d}\} \\
&= (U-d)\{a(\frac{db}{U-d} - a)\} \\
&\leq 0.
\end{aligned} \tag{C.40}$$

We conclude

$$\frac{d}{dx(0)}J_2^*(x(0)) < 0, \quad \text{for } x(0) < -(U-d)(a - \frac{1}{2}\frac{db}{U-d}).$$

Case H2: $x(0) \in (-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0)$.

According to Equation (4.187), we derive in the same way as in Equation (C.31)

$$\frac{d}{dx(0)}J_2^*(x(0)) = -\frac{x^2(0)}{U-d}, \tag{C.41}$$

and conclude

$$\frac{d}{dx(0)}J_2^*(x(0)) < 0, \quad \text{for } x(0) \in (-(U-d)(a - \frac{1}{2}\frac{db}{U-d}), 0).$$

Case H3: $x(0) \in (0, d(a - \frac{1}{2}\frac{db}{U-d}))$. According to Equation (4.187), we derive in the same way as in Equation (C.32)

$$\frac{d}{dx(0)}J_2^*(x(0)) = \frac{x^2(0)}{d}, \tag{C.42}$$

and conclude

$$\frac{d}{dx(0)}J_2^*(x(0)) > 0, \quad \text{for } x(0) \in (0, d(a - \frac{1}{2}\frac{db}{U-d})).$$

Case H4: $x(0) \in (d(a - \frac{1}{2}\frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b)$.

Since we have $S_2(x(0)) \leq \frac{1}{2}\frac{db}{U-d}$ for $x(0) \in (d(a - \frac{1}{2}\frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b)$, we can show in a similar way as in Case G4 in Section C.7 that the derivative of $J_2^*(x(0))$ with respect to $x(0)$ during the interval $(d(a - \frac{1}{2}\frac{db}{U-d}), da + (U - \sqrt{U(U-d)})b)$ is positive.

Case H5: $x(0) > da + (U - \sqrt{U(U-d)})b$.

According to Equation (4.187), we have for $x(0) > da + (U - \sqrt{U(U-d)})b$,

$$\frac{d}{dx(0)}J_2^*(x(0))$$

$$\begin{aligned}
&= \int_0^{a+b} 2(x(0) - dt)dt + \frac{d}{dx(0)} J_1^*(x(0) - d(a+b)) \\
&= 2x(0)(a+b) - d(a+b)^2 + \frac{d}{dx(0)} J_1^*(x(0) - d(a+b)) \quad (C.43) \\
&\geq 2(da + (U - \sqrt{U(U-d)})b)(a+b) - d(a+b)^2 \\
&= 2(da + (U - \sqrt{U(U-d)})b)(a+b) - d(a+b)^2 \\
&= 2da^2 + 2dab + 2(U - \sqrt{U(U-d)})ab \\
&\quad + 2(U - \sqrt{U(U-d)})b^2 - da^2 - 2dab - db^2 \\
&= da^2 + 2(U - \sqrt{U(U-d)})ab + 2(U - \sqrt{U(U-d)})b^2 - db^2, \\
&= da^2 + 2(U - \sqrt{U(U-d)})ab + (Ub^2 - 2\sqrt{U(U-d)}b^2 + (U-d)b^2), \\
&= da^2 + 2(U - \sqrt{U(U-d)})ab + (\sqrt{U} - \sqrt{U-d})^2 b^2, \\
&> 0. \quad (C.44)
\end{aligned}$$

Finally, we observe

$$\lim_{x(0) \rightarrow 0, x(0) > 0} \frac{d}{dx(0)} J_2^*(x(0)) = \lim_{x(0) \rightarrow 0, x(0) < 0} \frac{d}{dx(0)} J_2^*(x(0)) = 0.$$

In summary, we have validated Equation (4.188):

$$\begin{aligned}
\frac{d}{dx(0)} J_2^*(x_0) &< 0, & \text{for } x(0) < 0, \\
\frac{d}{dx(0)} J_2^*(x_0) &= 0, & \text{for } x(0) = 0, \\
\frac{d}{dx(0)} J_2^*(x_0) &> 0, & \text{for } x(0) > 0.
\end{aligned}$$

C.9 An Initial Production Control for Problem P_2

This is a modification of the Case **Case A3-2** for Problem P_2 . It also applies to Problem P_N for $N \geq 3$. In this case, we have $x(0) = 0$, $u(t) = U$ for $t \in [0, a)$, and $d = \frac{a}{a+b}U$. For $\Delta t \geq 0$, we define a class of controls $u_{\Delta t}(t)$ by

$$u_{\Delta t}(t) = \begin{cases} d, & \text{if } t \in [0, \Delta t), \\ U, & \text{if } t \in [\Delta t, a), \end{cases}$$

which leads to a trajectory with

$$x_{\Delta t}(t) = \begin{cases} 0, & \text{for } t \in [0, \Delta t], \\ (U-d)(t-\Delta t), & \text{for } t \in [\Delta t, a], \\ (U-d)(a-\Delta t) - d(t-a), & \text{for } t \in [a, a+b]. \end{cases}$$

Notice that $x(t)$ is identical to $x_{\Delta t}(t)$ for $\Delta t = 0$. Since we have $x(a) = (U-d)a = db$ and $x_{\Delta t}(a) = db - (U-d)\Delta t < db$ for $\Delta t > 0$, we cannot apply Property (4.118) for Problem P_2 . We directly use the definition of the performance measure and obtain

$$\begin{aligned} J_2(u_{\Delta t}) &= \int_{\Delta t}^a \{(U-d)(t-\Delta t)\}^2 dt + \int_a^{a+b} \{(U-d)(a-\Delta t) - d(t-a)\}^2 dt \\ &\quad + J_1^*((U-d)(a-\Delta t) - db) \\ &= \int_0^{a-\Delta t} \{(U-d)t\}^2 dt + \int_0^b \{(U-d)(a-\Delta t) - dt\}^2 dt \\ &\quad + J_1^*(-(U-d)\Delta t). \end{aligned} \quad (C.45)$$

We take the derivative with respect to Δt and obtain for $\Delta t \in (0, a - C_r)$,

$$\begin{aligned} \frac{d}{d\Delta t} J_2(u_{\Delta t}) &= (-1)\{(U-d)(a-\Delta t)\}^2 + \int_0^b -2(U-d)\{(U-d)(a-\Delta t) - dt\} dt \\ &\quad + \frac{d}{d\Delta t} J_1^*(-(U-d)\Delta t) \\ &= -(U-d)^2(a-\Delta t)^2 - 2(U-d)^2(a-\Delta t)b + (U-d)db^2 \\ &\quad - (U-d)\left(-\frac{(U-d)\Delta t^2}{U-d}\right) \\ &= -(U-d)^2\{(a-\Delta t)^2 + 2(a-\Delta t)b - \frac{d}{U-d}b^2\} - (U-d)^2(\Delta t)^2 \\ &= -(U-d)^2\{(a-\Delta t+b)^2 - \frac{U}{U-d}b^2 + (\Delta t)^2\}. \end{aligned} \quad (C.46)$$

From $d = \frac{a}{a+b}U$, we obtain $b = \frac{U-d}{d}a$. Replacing b in Equation (C.46) results in

$$\begin{aligned} \frac{d}{d\Delta t} J_2(u_{\Delta t}) &= -(U-d)^2\{(a-\Delta t+b)^2 - \frac{U}{U-d}b^2 + (\Delta t)^2\} \\ &= -(U-d)^2\left\{\left(\frac{U}{d}a - \Delta t\right)^2 - \frac{U}{U-d}\left\{\frac{U-d}{d}a\right\}^2 + (\Delta t)^2\right\} \\ &= -(U-d)^2\left\{\frac{U^2}{d^2}\left(a - \frac{d}{U}\Delta t\right)^2 - \frac{U(U-d)}{d^2}\{a\}^2 + (\Delta t)^2\right\} \\ &= -(U-d)^2\left\{\left(\frac{U^2}{d^2} - \frac{U(U-d)}{d^2}\right)a^2 - 2\frac{U}{d}a\Delta t + (\Delta t)^2 + (\Delta t)^2\right\} \end{aligned}$$

$$\begin{aligned}
&= -(U-d)^2 \left\{ \frac{U}{d} a^2 - 2 \frac{U}{d} a \Delta t + (\Delta t)^2 + (\Delta t)^2 \right\} \\
&= -(U-d)^2 \left\{ \frac{U}{d} a(a-2\Delta t) + 2(\Delta t)^2 \right\}
\end{aligned}$$

That means that for $\Delta t \in (0, \min(\frac{a}{2}, a - C_r))$, we have $\frac{d}{d\Delta t} J_2(u_{\Delta t}) < 0$ and consequently, the policy $u(t)$ is not optimal.

For Problem P_N ($N \geq 3$), the term $J_1^*(-(U-d)\Delta t)$ in Equation (C.45) is replaced by $J_{N-1}^*(-(U-d)\Delta t)$. Since the derivative of $J_{N-1}^*(-(U-d)\Delta t)$ with respect to Δt is $-(U-d)^2(\Delta t)^2$ for sufficiently small $\Delta t > 0$ and for all $N \geq 3$, we will have the same expression for $\frac{d}{d\Delta t} J(u_{\Delta t})$ in Equation (C.46). Thus, the Case A3-2 is validated in the same way for all $N \geq 2$.

C.10 Rewriting a quadratic equation (Case C4-2)

According to Equation (4.172), we have

$$\begin{aligned}
\frac{d}{dS} J_2(S) &= U(U+d) \left\{ S^2 + 2 \frac{x(0) - da + Ub}{U+d} S - \frac{2b}{U+d} \left(d(a + \frac{b}{2}) - x(0) \right) \right\} \\
&\quad - \frac{U}{U-d} (x(0) + US - d(a+b))^2.
\end{aligned} \tag{C.47}$$

Our goal in this section is to simplify this quadratic equation and solve for its roots. We rewrite this equation as follows:

$$\begin{aligned}
&\frac{d}{dS} J_2(S) \\
&= U(U+d) \left\{ S^2 + 2 \frac{x(0) - da + Ub}{U+d} S - \frac{2b}{U+d} \left(d(a + \frac{b}{2}) - x(0) \right) \right\} \\
&\quad - \frac{U}{U-d} (x(0) + US - d(a+b))^2 \\
&= U(U+d) S^2 + 2U(x(0) - da + Ub) S + 2Ub(x(0) - d(a + \frac{b}{2})) \\
&\quad - \frac{U}{U-d} (U^2 S^2 + 2(x(0) - d(a+b))US + (x(0) - d(a+b))^2) \\
&= (U(U+d) - \frac{U^3}{U-d}) S^2 + 2(U(x(0) - da + Ub) - \frac{U^2}{U-d} (x(0) - d(a+b))) S \\
&\quad + 2U(x(0) - d(a + \frac{b}{2}))b - \frac{U}{U-d} (x(0) - d(a+b))^2.
\end{aligned} \tag{C.48}$$

We now rewrite the coefficients. The coefficient of S^2 satisfies

$$U(U+d) - \frac{U^3}{U-d} = \frac{U(U^2 - d^2) - U^3}{U-d} = -\frac{Ud^2}{U-d}, \tag{C.49}$$

the coefficient of $S^1 = S$ satisfies

$$\begin{aligned}
 & 2(U(x(0) - da + Ub) - \frac{U^2}{U-d}(x(0) - d(a+b))) \\
 = & 2 \frac{U(x(0) - da + Ub)(U-d) - U^2(x(0) - d(a+b))}{U-d} \\
 = & 2 \frac{U^2(x(0) - da + Ub) - Ud(x(0) - da + Ub) - U^2(x(0) - d(a+b))}{U-d} \\
 = & 2 \frac{U^2(Ub) - Ud(x(0) - da)}{U-d} \\
 = & \frac{U}{U-d} 2(U^2b - d(x(0) - da)), \tag{C.50}
 \end{aligned}$$

and the coefficient of S^0 (the constant in Equation (C.48)) satisfies

$$\begin{aligned}
 & 2U(x(0) - d(a + \frac{b}{2}))b - \frac{U}{U-d}(x(0) - d(a+b))^2 \\
 = & 2Ubx(0) - 2Udab - Udb^2 - \frac{U}{U-d}(x^2(0) - 2x(0)d(a+b) + d^2(a+b)^2) \\
 = & x^2(0)\{-\frac{U}{U-d}\} + x(0)\{2(Ub + \frac{U}{U-d}d(a+b))\} \\
 & + \{-2Udab - Udb^2 - \frac{U}{U-d}d^2(a+b)^2\}. \tag{C.51}
 \end{aligned}$$

We simplify Equation (C.51) further by rewriting the coefficients of $x(0)$ and $x^0(0)$. We have

$$\{2(Ub + \frac{U}{U-d}d(a+b))\} = \{2 \frac{U(U-d)b + Ud(a+b)}{U-d}\} = \frac{U}{U-d} 2(Ub + da), \tag{C.52}$$

and

$$\begin{aligned}
 & \{-2Udab - Udb^2 - \frac{U}{U-d}d^2(a+b)^2\} \\
 = & \{ \frac{(-2Udab - Udb^2)(U-d) - Ud^2(a+b)^2}{U-d} \} \\
 = & \{ \frac{U}{U-d} \{ (-2dab - db^2)(U-d) - d^2(a+b)^2 \} \} \\
 = & \{ \frac{U}{U-d} \{-2Udab - Udb^2 + 2d^2ab + d^2b^2 - d^2a^2 - 2d^2ab - d^2b^2\} \} \\
 = & \{ \frac{U}{U-d} \{-2Udab - Udb^2 - d^2a^2\} \} \\
 = & \{ -\frac{U}{U-d} d \{ da^2 + 2Uab + Ub^2 \} \}. \tag{C.53}
 \end{aligned}$$

We now substitute Equations (C.52) and (C.53) into Equation (C.51) and obtain

$$x^2(0)\{-\frac{U}{U-d}\} + x(0)\{2(Ub + \frac{U}{U-d}d(a+b))\}$$

$$\begin{aligned}
& + \{-2Udab - Udb^2 - \frac{U}{U-d}d^2(a+b)^2\} \\
= & -\frac{U}{U-d}x^2(0) + \frac{U}{U-d}2(Ub+da)x(0) - \frac{U}{U-d}d\{da^2 + 2Uab + Ub^2\} \\
= & -\frac{U}{U-d}\{x^2(0) - 2(Ub+da)x(0) + d(da^2 + 2Uab + Ub^2)\}. \tag{C.54}
\end{aligned}$$

We now substitute the results of Equations (C.49), (C.50), and (C.54) into Equation (C.48) and obtain

$$\begin{aligned}
& \frac{d}{dS}J_2(S) \\
= & (U(U+d) - \frac{U^3}{U-d})S^2 + 2(U(x(0) - da + Ub) - \frac{U^2}{U-d}(x(0) - d(a+b)))S \\
& + 2U(x(0) - d(a + \frac{b}{2}))b - \frac{U}{U-d}(x(0) - d(a+b))^2 \\
= & -\frac{Ud^2}{U-d}S^2 + 2\frac{U}{U-d}(U^2b - d(x(0) - da))S \\
& - \frac{U}{U-d}\{x^2(0) - 2(Ub+da)x(0) + d(da^2 + 2Uab + Ub^2)\} \\
= & -\frac{Ud^2}{U-d}\{S^2 - 2\frac{U^2b - d(x(0) - da)}{d^2}S \\
& + \frac{x^2(0) - 2(Ub+da)x(0) + d(da^2 + 2Uab + Ub^2)}{d^2}\}. \tag{C.55}
\end{aligned}$$

For simplification, we introduce $f(x(0))$, x'_2 , and x''_2 as follows:

$$f(x(0)) = \frac{x^2(0) - 2(Ub+da)x(0) + d(da^2 + 2Uab + Ub^2)}{d^2}, \tag{C.56}$$

$$x'_2 = Ub + da - \sqrt{U(U-d)}b = da + b(U - \sqrt{U(U-d)}), \tag{C.57}$$

$$x''_2 = Ub + da + \sqrt{U(U-d)}b. \tag{C.58}$$

We now consider the function $f(x(0))$. We have

$$f(x(0)) = \frac{(x(0) - (Ub+da))^2 - (Ub+da)^2 + d(da^2 + 2Uab + Ub^2)}{d^2}.$$

Since we have

$$\begin{aligned}
(Ub+da)^2 - d(da^2 + 2Uab + Ub^2) &= U^2b^2 + 2Udab + d^2a^2 - d^2a^2 - 2Udab - Udb^2 \\
&= U(U-d)b^2,
\end{aligned}$$

we conclude

$$f(x(0)) = 0 \quad \text{if and only if} \quad x(0) \in \{x'_2, x''_2\}, \tag{C.59}$$

and

$$f(x(0)) < 0 \quad \text{if and only if} \quad x(0) \in (x'_2, x''_2). \quad (\text{C.60})$$

We observe

$$(da + b(U - \sqrt{U(U-d)}), d(a+b)) = (x'_2, d(a+b)) \subset (x'_2, x''_2). \quad (\text{C.61})$$

From Equations (C.59), (C.60) and (C.61), we conclude

$$f(x(0)) < 0, \quad \text{for } x(0) \in (da + b(U - \sqrt{U(U-d)}), d(a+b)), \quad (\text{C.62})$$

$$f(x(0)) > 0, \quad \text{for } x(0) \in (d(a - \frac{1}{2} \frac{db}{U-d}), da + b(U - \sqrt{U(U-d)})). \quad (\text{C.63})$$

We further define

$$S'_1 = \frac{U^2b - d(x(0) - da)}{d^2} - \sqrt{\left(\frac{U^2b - d(x(0) - da)}{d^2}\right)^2 - f(x(0))}, \quad (\text{C.64})$$

$$S'_2 = \frac{U^2b - d(x(0) - da)}{d^2} + \sqrt{\left(\frac{U^2b - d(x(0) - da)}{d^2}\right)^2 - f(x(0))}. \quad (\text{C.65})$$

According to Equation (C.55), we have for the second derivative

$$\frac{d^2}{dS^2} J_2(S) = -2 \frac{U d^2}{U-d} \left\{ S - \frac{U^2b - d(x(0) - da)}{d^2} \right\}. \quad (\text{C.66})$$

We evaluate Equation (C.66) at $S = S'_2$ and $S = S''_2$ and obtain

$$\begin{aligned} \frac{d^2}{dS^2} J_2(S'_2) &> 0, \\ \frac{d^2}{dS^2} J_2(S''_2) &< 0. \end{aligned}$$

Since our objective is to minimize, we focus on S'_2 . We rename S'_2 as $S_2(x(0))$ and have

$$S_2(x(0)) = \frac{U^2b - d(x(0) - da)}{d^2} - \sqrt{\left(\frac{U^2b - d(x(0) - da)}{d^2}\right)^2 - f(x(0))}. \quad (\text{C.67})$$

According to Equation (C.67), we conclude

$$S_2(x(0)) < 0, \quad \text{if and only if } f(x(0)) < 0. \quad (\text{C.68})$$

We now use Equations (C.62) and (C.68) to obtain

$$\begin{aligned} S_2(x(0)) &< 0, \quad \text{for } x(0) \in (da + b(U - \sqrt{U(U-d)}), d(a+b)), \\ S_2(x(0)) &> 0, \quad \text{for } x(0) \in (d(a - \frac{1}{2} \frac{db}{U-d}), da + b(U - \sqrt{U(U-d)})). \end{aligned} \quad (\text{C.69})$$

C.11 $S_2(x(0))$ is monotone decreasing (C4-2-2)

In this section, we show that the function $S_2(x(0))$ is monotone decreasing for $x(0) \in [d(a - C_r), d(a + \frac{b}{2})]$, i.e.,

$$\frac{d}{dx(0)} S_2(x(0)) < 0, \quad \text{for } x(0) \in (d(a - C_r), d(a + \frac{b}{2})). \quad (\text{C.70})$$

According to Equation (4.174), we obtain the following expression for the derivative of $S_2(x(0))$ with respect to $x(0)$:

$$\begin{aligned} \frac{d}{dx(0)} S_2(x(0)) &= -\frac{1}{d} - \frac{1}{2\sqrt{(\frac{U^2 b - d(x(0) - da)}{d^2})^2 - f(x(0))}} \left\{ 2\left(\frac{U^2 b - d(x(0) - da)}{d^2}\right)\left(-\frac{1}{d}\right) \right. \\ &\quad \left. - \frac{d}{dx(0)} f(x(0)) \right\}, \end{aligned} \quad (\text{C.71})$$

We now evaluate the term $\{2(\frac{U^2 b - d(x(0) - da)}{d^2})(-\frac{1}{d}) - \frac{d}{dx(0)} f(x(0))\}$:

$$\begin{aligned} &= 2\left(\frac{U^2 b - d(x(0) - da)}{d^2}\right)\left(-\frac{1}{d}\right) - \frac{d}{dx(0)} f(x(0)) \\ &= 2\left(\frac{U^2 b - d(x(0) - da)}{d^2}\right)\left(-\frac{1}{d}\right) - \frac{2x(0) - 2(Ub + da)}{d^2} \\ &= -\frac{2}{d^2}\left(\frac{U^2}{d}b - (x(0) - da) + x(0) - (Ub + da)\right) \\ &= -\frac{2}{d^2}\left(\frac{U^2}{d} - U\right)b \end{aligned} \quad (\text{C.72})$$

Substituting Equation (C.72) into Equation (C.71), we obtain

$$\begin{aligned} &\frac{d}{dx(0)} S_2(x(0)) \\ &= -\frac{1}{d} - \frac{1}{2\sqrt{(\frac{U^2 b - d(x(0) - da)}{d^2})^2 - f(x(0))}} \left(-\frac{2}{d^2}\right)\left(\frac{U^2}{d} - U\right)b \\ &= -\frac{1}{d} + \frac{(\frac{U^2}{d} - U)b}{d^2\sqrt{(\frac{U^2 b - d(x(0) - da)}{d^2})^2 - f(x(0))}} \\ &= \frac{(\frac{U^2}{d} - U)b - d\sqrt{(\frac{U^2 b - d(x(0) - da)}{d^2})^2 - f(x(0))}}{d^2\sqrt{(\frac{U^2 b - d(x(0) - da)}{d^2})^2 - f(x(0))}}. \end{aligned} \quad (\text{C.73})$$

Thus, we have

$$\frac{d}{dx(0)} S_2(x(0)) < 0$$

if and only if

$$\left(\frac{U^2}{d} - U\right)^2 b^2 - d^2 \left\{ \left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - f(x(0)) \right\} < 0. \quad (C.74)$$

According to Equation (C.56), we now evaluate the expression $\left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - f(x(0))$:

$$\begin{aligned} & \left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - f(x(0)) \\ = & \left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - \frac{x^2(0) - 2(Ub + da)x(0) + d(da^2 + 2Uab + Ub^2)}{d^2} \\ = & \frac{1}{d^2} \left(\left(\frac{U^2}{d} b - (x(0) - da) \right)^2 - x^2(0) + 2(Ub + da)x(0) - d(da^2 + 2Uab + Ub^2) \right) \\ = & \frac{1}{d^2} \left(\frac{U^4}{d^2} b^2 - 2 \frac{U^2}{d} b(x(0) - da) + (x(0) - da)^2 \right. \\ & \quad \left. - (x(0) - da)^2 + 2(Ub)x(0) - d(2Uab + Ub^2) \right) \\ = & \frac{1}{d^2} \left(\frac{U^4}{d^2} b^2 - 2 \frac{U^2}{d} b(x(0) - da) + 2(Ub)x(0) - d(2Uab + Ub^2) \right) \\ = & \frac{1}{d^2} \left(\left(\frac{U^4}{d^2} - U d \right) b^2 - 2 \left(\frac{U^2}{d} b - Ub \right) (x(0) - da) \right) \end{aligned} \quad (C.75)$$

Substituting Equation (C.75) in Equation (C.74), we obtain

$$\begin{aligned} & \left(\frac{U^2}{d} - U \right)^2 b^2 - d^2 \left\{ \left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - f(x(0)) \right\} \\ = & \left(\frac{U^2}{d} - U \right)^2 b^2 - \left(\left(\frac{U^4}{d^2} - U d \right) b^2 - 2 \left(\frac{U^2}{d} b - Ub \right) (x(0) - da) \right) \\ = & 2 \left(\frac{U^2}{d} b - Ub \right) (x(0) - da) - \left(\left(\frac{U^4}{d^2} - U d \right) b^2 - \left(\frac{U^2}{d} - U \right)^2 b^2 \right) \\ = & 2 \left(\frac{U^2}{d} b - Ub \right) (x(0) - da) - \left(\left(\frac{U^4}{d^2} - U d \right) - \left(\frac{U^2}{d} - U \right)^2 \right) b^2. \end{aligned} \quad (C.76)$$

We rewrite $\left(\frac{U^2}{d} b - Ub \right)$ and obtain

$$\left(\frac{U^2}{d} b - Ub \right) = \frac{U^2 - U d}{d} b = \frac{U - d}{d} Ub. \quad (C.77)$$

For $\left(\left(\frac{U^4}{d^2} - U d \right) - \left(\frac{U^2}{d} - U \right)^2 \right)$, we obtain

$$\begin{aligned} \left(\left(\frac{U^4}{d^2} - U d \right) - \left(\frac{U^2}{d} - U \right)^2 \right) &= \frac{U^4}{d^2} - U d - \frac{U^4}{d^2} + 2 \frac{U^3}{d} - U^2 \\ &= -U d + 2 \frac{U^3}{d} - U^2 \\ &= U \frac{2U^2 - d^2 - U d}{d} \\ &= U \frac{U^2 - d^2 + U^2 - U d}{d} \end{aligned}$$

$$\begin{aligned}
&= U \frac{(U-d)(U+d) + (U-d)U}{d} \\
&= \frac{U-d}{d} U((U+d) + U). \tag{C.78}
\end{aligned}$$

Substituting Equations (C.77) and (C.78) into Equation (C.76), we obtain

$$\begin{aligned}
&2 \frac{U-d}{d} U b(x(0) - da) - \frac{U-d}{d} U((U+d) + U) b^2 \\
&= 2 \left(\frac{U-d}{d} U b \right) (x(0) - da) - \frac{1}{2} ((U+d) + U) b^2 \\
&= 2 \left(\frac{U-d}{d} U b \right) (x(0) - (U b + d(a + \frac{b}{2}))).
\end{aligned}$$

We thus conclude

$$\frac{d}{dx(0)} S_2(x(0)) < 0, \quad \text{for } x(0) < U b + d(a + \frac{b}{2}),$$

which validates Equation (C.70). Thus, the maximum value of $S_2(x(0))$ during the interval $[d(a - \frac{1}{2} \frac{db}{U-d}), d(a + \frac{b}{2})]$ is obtained at $x(0) = a - \frac{1}{2} \frac{db}{U-d}$.

C.12 Evaluation of $S_2(d(a - \frac{1}{2} \frac{db}{U-d}))$ (Case C4-2-2)

In this section, we show

$$S_2(d(a - \frac{1}{2} \frac{db}{U-d})) = a - \frac{1}{2} \frac{db}{U-d}. \tag{C.79}$$

According to Equation (4.174), $S_2(x(0))$ is given by

$$S_2(x(0)) = \frac{U^2 b - d(x(0) - da)}{d^2} - \sqrt{\left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - f(x(0))}. \tag{C.80}$$

where $f(x(0)) = \frac{x^2(0) - 2(Ub+da)x(0) + d(da^2 + 2Uab + Ub^2)}{d^2}$. We first rewrite $f(x(0))$. We have

$$\begin{aligned}
f(x(0)) &= \frac{1}{d^2} \{x^2(0) - 2(Ub + da)x(0) + d^2 a^2 + 2Udab + Udb^2\} \\
&= \frac{1}{d^2} \{(x(0) - (Ub + da))^2 - (U^2 b^2 + 2Udab + d^2 a^2) + d^2 a^2 + 2Udab + Udb^2\} \\
&= \frac{1}{d^2} \{(x(0) - (Ub + da))^2 - U(U - d)b^2\}.
\end{aligned}$$

We now evaluate the function $f(x(0))$ for $x(0) = d(a - \frac{1}{2} \frac{db}{U-d})$:

$$f(d(a - \frac{1}{2} \frac{db}{U-d}))$$

$$\begin{aligned}
&= \frac{1}{d^2} \left\{ \left(d \left(a - \frac{1}{2} \frac{db}{U-d} \right) - (Ub + da) \right)^2 - U(U-d)b^2 \right\} \\
&= \frac{1}{d^2} \left\{ \left(-d \frac{1}{2} \frac{db}{U-d} - Ub \right)^2 - U(U-d)b^2 \right\} \\
&= \frac{b^2}{(U-d)^2 d^2} \left\{ \left(\frac{d^2}{2} + U(U-d) \right)^2 - U(U-d)^3 \right\} \\
&= \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{d^4}{4} + U(U-d)d^2 + U^2(U-d)^2 - U(U-d)^3 \right\} \\
&= \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{d^4}{4} + U(U-d)d^2 + U(U-d)^2 d \right\} \\
&= \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{d^4}{4} + U^2(U-d)d \right\}.
\end{aligned}$$

We proceed to evaluate the expression $\left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - f(x(0))$ under the squareroot (see Equation (C.80)) for $x(0) = d(a - \frac{1}{2} \frac{db}{U-d})$:

$$\begin{aligned}
&\left(\frac{U^2 b - d(x(0) - da)}{d^2} \right)^2 - f\left(d\left(a - \frac{1}{2} \frac{db}{U-d}\right)\right) \\
&= \left(\frac{U^2 b - d\left(d\left(a - \frac{1}{2} \frac{db}{U-d}\right) - da\right)}{d^2} \right)^2 - \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{d^4}{4} + U^2(U-d)d \right\} \\
&= \left(\frac{U^2 b + d^2 \frac{1}{2} \frac{db}{U-d}}{d^2} \right)^2 - \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{d^4}{4} + U^2(U-d)d \right\} \\
&= \left(\frac{U^2 b}{d^2} + \frac{1}{2} \frac{db}{U-d} \right)^2 - \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{d^4}{4} + U^2(U-d)d \right\} \\
&= \frac{U^4}{d^4} b^2 + \frac{U^2}{d^2} b \frac{db}{U-d} + \frac{1}{4} \frac{d^2 b^2}{(U-d)^2} - \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{d^4}{4} + U^2(U-d)d \right\} \\
&= \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{U^4(U-d)^2}{d^2} + U^2(U-d)d + \frac{d^4}{4} - \left(\frac{d^4}{4} + U^2(U-d)d \right) \right\} \\
&= \frac{b^2}{(U-d)^2 d^2} \left\{ \frac{U^4(U-d)^2}{d^2} \right\} \\
&= \frac{b^2}{d^2} \left\{ \frac{U^4}{d^2} \right\} \\
&= \left(\frac{U^2}{d^2} b \right)^2. \tag{C.81}
\end{aligned}$$

Therefore, we have according to Equations (C.80) and (C.81)

$$\begin{aligned}
S_2\left(d\left(a - \frac{1}{2} \frac{db}{U-d}\right)\right) &= \frac{U^2 b - d\left(d\left(a - \frac{1}{2} \frac{db}{U-d}\right) - da\right)}{d^2} \\
&\quad - \left\{ \left(\frac{U^2 b - d\left(d\left(a - \frac{1}{2} \frac{db}{U-d}\right) - da\right)}{d^2} \right)^2 \right. \\
&\quad \left. - f\left(d\left(a - \frac{1}{2} \frac{db}{U-d}\right)\right) \right\}^{\frac{1}{2}} \\
&= \frac{U^2 b - d\left(d\left(a - \frac{1}{2} \frac{db}{U-d}\right) - da\right)}{d^2} - \sqrt{\left(\frac{U^2}{d^2} b \right)^2}
\end{aligned}$$

$$= \frac{1}{2} \frac{db}{U-d}, \tag{C.82}$$

which validates Equation (C.80).

REFERENCES

- [1] Abdul-Razaq, T. and C. Potts. "Dynamic Programming State-Space Relaxation for Single-Machine Scheduling," *Journal of the Operational Research Society*, 39, pp. 141-152, 1988.
- [2] Ahmadi, R. and U. Bagchi. "Single Machine Scheduling to Minimize Equalness Subject to Deadlines," Working Paper 86/86-4-17, Department of Management, University of Texas, Austin, 1986.
- [3] Ahmadi, R. and U. Bagchi. "Just-in-Time Scheduling in Single Machine Systems," Working Paper 85/86-4-21, Department of Management, University of Texas, Austin, 1986.
- [4] Akella, R. A., Y. F. Choong, and S. B. Gershwin. "Performance of Hierarchical Production Scheduling Policy," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, CHMT-7:3, pp. 225-240, 1984.
- [5] Akella, R. A. and P. R. Kumar. "Optimal Control of Production Rate in a Failure Prone Manufacturing System," *IEEE Transactions on Automatic Control*, AC-31:2, pp. 116-126, 1986.
- [6] Anger, F.D., C.-Y. Lee, and L. A. Martin-Vega, "Single Machine Scheduling with Tight Windows," Research Report 86-16, Department of Industrial and Systems Engineering, University of Florida, 1986.
- [7] Bagchi, U. "Scheduling to Minimize Earliness and Tardiness Penalties With a Common Due Date," Working Paper, Department of Management, University of Texas, Austin, 1985.
- [8] Bagchi, U. "Due Date or Deadline Assignment to Multi-Job Customer Orders," Working Paper, Department of Management, University of Texas, Austin, 1989.
- [9] Bagchi, U., Y. Chang and R. Sullivan. "Minimizing Absolute and Squared Deviations of Completion Times with Different Earliness and Tardiness Penalties and a Common Due Date," *Naval Research Logistics Quarterly*, 34, pp. 739-751, 1987.
- [10] Bagchi, U., R. Sullivan, and Y. Chang. "Minimizing Mean Absolute Deviation of Completion Times About a Common Due Date," *Naval Research Logistics Quarterly*, 33, pp. 227-240, 1986.
- [11] Bagchi, U., R. Sullivan, and Y. Chang. "Minimizing Mean Squared Deviation of Completion Times About a Common Due Date," *Management Science*, 33, pp. 894-906, 1987.
- [12] Bai, S. X. and J. H. Burhanpurwala. "Hierarchical Production Control for a Manufacturing System with Dynamic Setup Changes," Research Report 94-2, Department of Industrial and Systems Engineering, Gainesville, 1994.

- [13] Bai, S. X. and M. Elhafi. "Real-Time Scheduling of a Manufacturing System with non-resumable Setup Changes," Research Report 93-28, Department of Industrial and Systems Engineering, Gainesville, 1993.
- [14] Bai, S. X. and S. B. Gershwin. "Scheduling Manufacturing Systems with Work-In-Process Inventory Control: Single-Part-Type Systems," To appear in *IIE Transactions*, 1994.
- [15] Bai, S. X. and S. B. Gershwin. "Scheduling Manufacturing Systems with Work-In-Process Inventory Control: Multiple-Part-Type Systems," *International Journal of Production Research*, Vol. 32, No. 2, pp. 365-385, 1994.
- [16] Bai, S. X. and S. B. Gershwin. "Scheduling Manufacturing Systems with Work-In-Process Inventory Control: Reentrant Systems," To appear in *IIE Transactions*, 1994.
- [17] Baker, K. R. "Introduction to Sequencing and Scheduling," John Wiley and Sons, New York, 1974.
- [18] Baker, K. R. and A. Chadowitz. "Algorithms for Minimizing Earliness and Tardiness Penalties With a Common Due Date," Working Paper No. 240, Amos Tuck School of Business Administration, Dartmouth College, Hanover, N. H., 1989.
- [19] Baker, K. R. and G. D. Scudder. "On the Assignment of Optimal Due Dates," *Journal of the Operational Research Society*, 40, pp. 93-95, 1989.
- [20] Baker, K. R. and G. D. Scudder. "Sequencing with Earliness and Tardiness Penalties: A Review," *Operations Research*, 38:1, pp. 22-36, 1990.
- [21] Bector, C., Y. Gupta, and M. Gupta. "Determination of an Optimal Common Due Date and Optimal Sequence in a Single Machine Job Shop," *International Journal of Production Research*, 26, pp. 613-628, 1988.
- [22] Bryson, A. E., Jr. and Y.-C. Ho. "Applied Optimal Control: Optimization, Estimation, and Control," Blaisdell Publishing, Waltham, MA, 1969.
- [23] Chand, S. and H. Schneeberger. "Single Machine Scheduling to Minimize Weighted Earliness Subject to No Tardy Jobs," *European Journal of Operational Research*, 34, pp. 221-230, 1988.
- [24] Cheng, T. C. E. "Optimal Due Date Determination and Sequencing of n Jobs on a Single Machine," *Journal of the Operational Research Society*, 35, pp. 433-437, 1984.
- [25] Cheng, T. C. E. "An Algorithm for the CON Due Date Determination and Sequencing Problem," *Computers and Operations Research*, 14, pp.537-542, 1987.
- [26] Cheng, T. C. E. "Optimal Common Due Date with Limited Completion Time Deviation," *Computers and Operations Research*, 15, pp. 91-96, 1988.
- [27] Cheng, T. C. E. and M. Gupta. "Survey of Scheduling Research Involving Due Date Determination Decisions," *European Journal of Operational Research*, 38, pp. 156-166, 1989.
- [28] Cook, S. A. "The complexity of theorem-proving procedures," Proceedings of the 3rd Annual ACM Symposium on the Theory of Computing, pp. 151-158, 1971.
- [29] Coffman, E. G., Jr. (ed.). "Computer and Job Shop Scheduling Theory," Wiley, New York, 1976.

- [30] Conway, R. W., W. L. Maxwell, and L. W. Miller. "Theory of Scheduling," Addison-Wesley, Reading, MA, 1967.
- [31] Davis, J. and J. Kanet. "Single Machine Scheduling With a Nonregular Convex Performance Measure," Working Paper, Department of Management, Clemson University, Clemson, S.C., 1988.
- [32] De, P., J. Ghosh, and C. Wells. "A Note on the Minimization of Mean Squared Deviation of Completion Times About a Common Due Date." *Management Science*, 35, pp. 1143-1147, 1989.
- [33] De, P., J. Ghosh, and C. Wells. "Scheduling About a Common Due Date With Earliness and Tardiness Penalties," Working Paper, University of Dayton, Dayton OH, 1989.
- [34] Dempster, M. A. H., J. K. Lenstra, and A. H. G. Rinnooy Kan (eds.). "Deterministic and Stochastic Scheduling," Vol. 1, Reidel, Dordrecht, Netherlands, 1982.
- [35] Eilon, S. and I. Chowdhury. "Minimizing Waiting Time Variance in the Single Machine Problem," *Management Science*, 23, pp. 567-575, 1977.
- [36] Elsayed, E. A. and C. C. Hwang. "Analysis of Two Stage Manufacturing Systems with Buffer Storage and Redundant Machines," *International Journal of Production Research*, 24, pp. 187, 1986.
- [37] Emmons, H. "One-machine sequencing to minimize certain functions of job tardiness," *Operations Research*, 17, pp. 701-715, 1969.
- [38] Emmons, H. "Scheduling to a Common Due Date on Parallel Uniform Processors," *Naval Research Logistics Quarterly*, 34, 803-810, 1987.
- [39] Federgruen, A. and G. Mosheiov. "Efficient Algorithms for Scheduling Problems with General Earliness and Tardiness Cost Structures," Working Paper, Graduate School of Business, Columbia University, 1991.
- [40] French, S. "Sequencing and Scheduling: An Introduction to the Mathematics of the Job-Shop," John Wiley and Sons, 1982.
- [41] Fry, T. and G. Leong. "A Bi-Criterion Approach to Minimizing Inventory Costs on a Single Machine When Early Shipments are Forbidden," *Computers and Operations Research*, 14, pp. 363-368, 1987.
- [42] Fry, T., R. Armstrong, and J. Blackstone. "Minimizing Weighted Absolute Deviation in Single Machine Scheduling," *IEEE Transactions*, 19, pp. 445-450, 1987.
- [43] Fry, T., K. Darby-Dowman, and R. Armstrong. "Single Machine Scheduling to Minimize Mean Absolute Lateness," Working Paper, College of Business Administration, University of South Carolina, CO, 1988.
- [44] Fry, T., G. Leong, and T. Rakes. "Single Machine Scheduling: A Comparison of Two Solution Procedures," *OMEGA* 15, pp. 277-282, 1987.
- [45] Garey, M. R. and D. S. Johnson. "Strong NP-completeness results: motivation, examples and implications," *Journal of the Association of Computing Machines*, 25, pp. 499-508, 1978.
- [46] Garey, M. R. and D. S. Johnson. "Computers and Intractability: A Guide to the Theory of NP-completeness," Freeman, San Francisco, 1979.

- [47] Garey, M. R., R. Tarjan, and G. Wilfong. "One-Processor Scheduling With Symmetric Earliness and Tardiness Penalties," *Mathematics and Operations Research*, 13, pp. 330-348, 1988.
- [48] Gershwin, S. B. "Hierarchical flow control: a framework for scheduling and planning discrete events in manufacturing systems," *Proceedings of the IEEE, Special Issue on Dynamics of Discrete Event Systems*, 77:1, pp. 195-209, 1989.
- [49] Gershwin, S. B., R. A. Akella, and Y. F. Choong. "Short-term production scheduling of an automated manufacturing facility," *IBM Journal of Research and Development*, 29:4, pp. 392-400, 1985.
- [50] Graham, R. L., E. L. Lawler, J. K. Lenstra, and A. H. G. Rinnooy Kan. "Optimization and approximation in deterministic sequencing and scheduling: a survey," *Annals of Discrete Mathematics*, 5, pp. 287-326, 1979.
- [51] Graves, S. C. "A review of production scheduling," *Operations Research*, 26, pp. 200-203, 1978.
- [52] Gupta, S. and J. Kyparisis. "Single Machine Scheduling Research," *OMEGA*, 15, pp. 207-227, 1987.
- [53] Gupta, S. and T. Sen. "Minimizing a Quadratic Function of Job Lateness on a Single Machine," *Engineering Costs and Production Economist*, 7, pp. 181-194, 1983.
- [54] Hall, N. G. "Single and Multi-Processor Models for Minimizing Completion Time Variance," *Naval Research Logistics Quarterly*, 33, pp. 49-54, 1986.
- [55] Hall, N. G. and M. E. Posner. "Earliness-Tardiness Scheduling Problems, I: Weighted Deviation of Completion Times about a Common Due Date," *Operations Research*, 39, pp. 836-846, 1991.
- [56] Hall, N. G., W. Kubiak, and S. P. Sethi. "Earliness-Tardiness Scheduling Problems, II: Deviation of Completion Times about a Restrictive Common Due Date," *Operations Research*, 39, pp. 847-856, 1991.
- [57] Hestenes, M. R. "Calculus of Variations and Optimal Control Theory," Wiley, New York, 1966.
- [58] Hoogeveen, J. A. and S. L. Van De Velde. "Scheduling Around a Small Common Due Date," *European Journal of Operational Research*, 55, pp. 237-242, 1991.
- [59] Hopp, W. J., N. Pati, and P. C. Jones. "Optimal Inventory Control in a Production Flow System with Failures," Technical Report 88-07, Department of Industrial Engineering and Management Sciences, Northwestern University, IL, 1988.
- [60] Hu, J.-Q. and D. Xiang. "Optimal Production Control for Failure Prone Manufacturing Systems with Deterministic Machine Up Times," Working Paper, Boston University, 1992.
- [61] Ignall, E. and A. Silver. "The Output of a Two Stage System with Unreliable Machines and Limited Storage," *A.I.E.E. Transactions*, 9, pp. 183 ff., 1977.
- [62] Johnson, D. S. "The NP-completeness column: an ongoing guide," *Journal of Algorithms*, 4, pp. 189-203, 1983.
- [63] Kanet, J. "Minimizing the Average Deviation of Job Completion Times About a Common Due Date," *Naval Research Logistics Quarterly*, 28, pp. 643-651, 1981.

- [64] Karp, R. M. "Reducibility among combinatorial problems," Complexity of Computer Computations, R. E. Miller and J. W. Thatcher (eds.), Plenum Press, New York, pp. 85-103, 1972.
- [65] Karp, R. M. "On the computational complexity of combinatorial problems," *Networks*, 5, pp. 45-68, 1975.
- [66] Kimemia, J. and S. B. Gershwin. "An Algorithm for the Computer Control of a Flexible Manufacturing System," *IEEE Transactions*, 15:4, pp. 353-362, 1983.
- [67] Kirk, D. E. "Optimal Control Theory, An Introduction," *Network Series*, Prentice-Hall, 1970.
- [68] Krämer, F.-J. "Essays in Production Scheduling with Work-In-Process related performance measures." Ph.D. Dissertation, University of Florida, Gainesville, FL, 1994.
- [69] Krämer, F.-J. and S. X. Bai. "A Production Flow Control Problem with Periodic Maintenance," Research Report, Industrial and Systems Engineering, University of Florida, Gainesville, 1994.
- [70] Krämer, F.-J. and S. X. Bai. "Optimal Control of a Production System with Periodic Maintenance: Part I: Solution of the one-cycle Problem," Working Paper, Industrial and Systems Engineering, University of Florida, Gainesville, 1994.
- [71] Krämer, F.-J. and S. X. Bai. "Optimal Control of a Production System with Periodic Maintenance: Part II: Solution of the N -cycle Problem," Working Paper, Industrial and Systems Engineering, University of Florida, Gainesville, 1994.
- [72] Krämer, F.-J. and C.-Y. Lee. "Common due window scheduling," to appear in *Production and Operations Management*, 1994.
- [73] Krämer, F.-J., and C.-Y. Lee. "Due Window Scheduling for Parallel Machines," to appear in: *Mathematical and Computer Modelling*, Special Issue on Scheduling: Theory and Applications, 1994.
- [74] Lakshminarayan, S., R. Lakshmanan, R. L. Papineau, and R. Rochette. "Optimal Single-Machine Scheduling with Earliness and Tardiness Penalties," *Operations Research*, 26, pp. 1079-1082, 1978.
- [75] Lawler, E. L., J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys. "Sequencing and Scheduling: Algorithms and Complexity," Report BS-R8909, Centre for Mathematics and Computer Science, Amsterdam, Netherlands, 1989.
- [76] Lee, C.-Y. "Earliness-Tardiness Scheduling Problems with Constant Size of Due-Date Window," Working Paper 91-17, Department of Industrial and Systems Engineering, University of Florida, 1991.
- [77] Lee, C.-Y., S. Danusaputro, and C.S. Lin. "Minimizing Weighted Number of Tardy Jobs and Weighted Earliness-Tardiness Penalties about a Common Due Date," *Computers and Operations Research*, 18, pp. 379-389, 1991.
- [78] Lenstra, J. K. and A. H. G. Rinnooy Kan. "New directions in scheduling theory," *Operations Research Letters*, 2, pp. 255-259, 1984.
- [79] Lenstra, J. K., A. H. G. Rinnooy Kan, and P. Brucker. "Complexity of machine scheduling problems," *Annals of Discrete Mathematics*, 1, pp. 343-362, 1977.

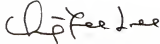
- [80] Liman, S. D. and C.-Y. Lee. "Error Bound for a Heuristic on the Common Due-Date Scheduling Problem," *ORSA Journal of Computing*, 5, 420-425, 1993.
- [81] Lou, S. X. C., S. P. Sethi, and Q. Zhang. "Optimal Feedback Production Planning in a Stochastic Two-Machine Flowshop," to appear in *European Journal of Operational Research*, Special Issue on Stochastic Control Theory and Operational Research, 1994.
- [82] Malathronas, J. P., J. D. Perkins, and R. L. Smith. "The Availability of a System of Two Unreliable Machines Connected by an Intermediate Storage Tank," *A.I.E.E. Transactions*, 15, pp. 195 ff., 1983.
- [83] Nahmias, S. "Production and Operations Analysis," Irwin Inc., Boston, MA, 1989.
- [84] Ow, P. and T. Morton. "The Single Machine Early-Tardy Problem," *Management Science*, 35, pp. 177-191, 1989.
- [85] Panwalkar, S., M. Smith, and A. Seidmann. "Common Due Date Assignment to Minimize Total Penalty for the One Machine Scheduling Problem," *Operations Research*, 30, pp. 391-399, 1982.
- [86] Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. "L. S. Pontryagin Selected Works," Vol. 4. The Mathematical Theory of Optimal Processes. Gordon and Breach Science Publishers, Montreux, Switzerland, 1986.
- [87] Presman, E., S. P. Sethi, and Q. Zhang. "Optimal Feedback Production Planning in a Stochastic N-Machine Flowshop," to appear in *Proceedings of the 12th World Congress of International Federation of Automatic Control*, Sydney, Australia, 1993.
- [88] Quaddus, M. "A Generalized Model of Optimal Due-Date Assignment by Linear Programming," *Journal of the Operational Research Society*, 38, pp. 353-359, 1987.
- [89] Raghavachari, M. "A V-Shape Property of Optimal Schedule of Jobs About a Common Due Date," *European Journal of Operational Research*, 23, pp. 401-402, 1986.
- [90] Raghavachari, M. "Scheduling Problems With Non-Regular Penalty Functions - A Review," *Opsearch*, 25, pp. 141-164, 1988.
- [91] Rinnooy Kan, A. H. G. "Machine Scheduling Problems: Classification, Complexity and Computations," Nijhoff, The Hague, Netherlands, 1976.
- [92] Ryzin, G. van, S. X. C. Lou, and S. B. Gershwin. "Production control for a two-machine system," to appear in *IIE Transactions*, 1994.
- [93] Seidmann, A., S. Panwalkar, and M. Smith. "Optimal Assignment of Due-Dates for a Single Processor Scheduling Problem," *International Journal of Production Research*, 19, pp. 393-399, 1981.
- [94] Sen, T. and S. Gupta. "A State-of-the-Art Survey of Static Scheduling Research Involving Due-Dates," *OMEGA*, 12:1, pp. 62-76, 1984.
- [95] Sethi, S. P., H. Yan, Q. Zhang, and X. Y. Zhou. "Feedback production planning in a stochastic two machine flow shop: Asymptotic analysis and computational results," *International Journal of Production Economics*, 30-31, pp. 79-93, 1993.
- [96] Sethi, S. P. and Q. Zhang. "Asymptotic optimality of hierarchical control in stochastic manufacturing systems: A Review," NSF Workshop on Hierarchical Control for Real-Time Scheduling of Manufacturing Systems, Lincoln, NH, 1992.

- [97] Sethi, S. P., Q. Zhang, and X. Y. Zhou. "Hierarchical Production Controls in a Stochastic Two-machine Flowshop with a Finite Internal Buffer," *Proceedings of the 31st Conference on Decision and Control*, Tucson, AZ, pp. 2074-2079, 1992.
- [98] Sethi, S. P., Q. Zhang, and X. Y. Zhou. "Hierarchical Control for Manufacturing Systems with Machines in Tandem," *Stochastics and Stochastics Reports*, 41, pp. 89-118, 1992.
- [99] Sethi, S. P. and X. Y. Zhou. "Feedback Controls in Two-Machine Flowshops, Part I: Optimal Controls in the Deterministic Case," Working Paper, Department of Management, University of Toronto, Canada, 1993.
- [100] Sethi, S. P. and X. Y. Zhou. "Feedback Controls in Two-Machine Flowshops, Part II: Asymptotic Optimal Controls in the Stochastic Case," Working Paper, Department of Management, University of Toronto, Canada, 1993.
- [101] Sharifnia, A., M. Caramanis, and S. B. Gershwin. "Dynamic Setup Scheduling and Flow Control in Manufacturing Systems," *Discrete Event Dynamic Systems: Theory and Applications*, Vol. 1, pp. 149-175, 1991.
- [102] Sidney, J. B. "Optimal Single-Machine Scheduling with Earliness and Tardiness Penalties," *Operations Research*, 25, pp. 62-69, 1977.
- [103] Srinivatsan, N. and S. B. Gershwin. "Selection of Setup Times in a Hierarchically Controlled Manufacturing System," *Proceedings of the 29th IEEE Conference on Decision and Control*, Hawaii, 1990.
- [104] Szwarc, W. "Single Machine Scheduling to Minimize Absolute Deviation of Completion times From a Common Due Date," *Naval Research Logistics Quarterly*, 36, pp. 663-673, 1989.
- [105] Ullman, J. D. "NP-Complete scheduling problems," *Journal of Computation System Sciences*, 10, pp. 384-393, 1975.
- [106] Ullman, J. D. "Complexity of sequencing problems," *Computer and Job Shop Scheduling Theory*, E. G. Coffman (ed.), Wiley, New York, pp. 139-164, 1976.
- [107] Vani, V. and M. Raghavachari. "Deterministic and Random Single Machine Scheduling With Variance Minimization," *Operations Research*, 35, pp. 111-120, 1987.
- [108] Wijngaard, J. "The Effect of Inter Stage Buffer Storage on the Output of Two Unreliable Production Units in Series With Different Production Rates," *A.I.E.E. Transactions*, 11, pp. 42 ff., 1979.
- [109] Yano, C. and Y. Kim. "Algorithms for a Class of Single-Machine Weighted Tardiness and Earliness Problems," *European Journal of Operational Research*, 51, pp. 1-12 1989.

BIOGRAPHICAL SKETCH

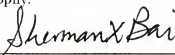
Franz-Josef Krämer was born in Daun, Germany, on July 8, 1965. He received the degree Diplom-Informatiker from the Department of Computer Science at the University of Kaiserslautern, Germany, in June 1989. In July 1989, he came to the United States to pursue his studies in the Department of Industrial and Systems Engineering. He received his Master of Science degree in this department in December 1990 and subsequently pursued the Ph.D. degree with specialization in operations research. Throughout his graduate studies he has been working as a graduate teaching and research assistant.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Chung-Yee Lee, Chairman
Associate Professor of Industrial
and Systems Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



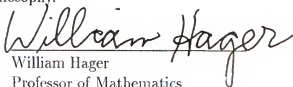
Sherman X. Bai, Cochairman
Assistant Professor of Industrial
and Systems Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



D. Jack Elzinga
Professor of Industrial
and Systems Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



William Hager
Professor of Mathematics

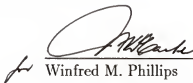
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Sencer Yeralan
Associate Professor of Industrial
and Systems Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1994



Winfred M. Phillips
Dean, College of Engineering

Karen A. Holbrook
Dean, Graduate School